Describing Cognitive Change in Normal Aging and Early-stage Dementia Using Measures of Verbal Fluency

A Thesis Submitted to
the College of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in the Department of Psychology
University of Saskatchewan
Saskatoon

By
Nicole Haugrud

PERMISSION TO USE

DISCLAIMER

Abstract

Verbal fluency tests require individuals to produce as many words as possible in a one minute trial either belonging to a specific category (semantic fluency) or starting with a specific letter (phonemic fluency). Researchers have proposed comparing subcomponents of fluency production, clustering (grouping semantically or phonemically related words) and switching (shifting between clusters; Abwender et al., 2001; Lanting et al., 2009; Troyer et al., 1997). The objective of the current research was to investigate measures of clustering and switching on verbal fluency tasks for healthy individuals and individuals diagnosed with dementia. Study 1 involved the development of a computer scoring program which was shown to produce more accurate and time efficient scoring. Study 2 compared clustering and switching variables across the healthy adult age span. The older age group produced fewer semantic fluency total words due to reduced hard switching, consistent with the frontal executive hypothesis of healthy aging (MacPherson et al., 2002). Study 3 compared healthy older adults to individuals diagnosed with AD. Measures of clustering and switching did not reliably differentiate AD from healthy aging, which could have resulted from the heterogeneity of the AD group. Study 4 compared clustering and switching variables longitudinally in an AD sample. When initial stage of symptom severity was controlled for, individuals at early stages of AD showed decline in phonemic total word production over time due to decline in switching ability and continued to show slight decline on semantic fluency over time, consistent with the progression of AD to prefrontal lobe regions (Levy & Chelune, 2007). The goal of study 5 was to determine which variables best differentiated subtypes of dementia. Using a homogeneous group of individuals diagnosed with AD, dementia subtypes showed differential patterns of clustering and switching impairment. Results from this body of research supports the use of the variables total word production, hard switches, and cluster switches on phonemic fluency, and the use of the variables total word production, average cluster size, hard switches, and cluster switches on semantic fluency.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Drs. Margaret Crossley and Mirna Vrbancic for their continued support and encouragement throughout my graduate training. Thank you to the faculty members and professional affiliates I have had the opportunity to work with and learn from over the past five years including individuals in the Department of Psychology, the Rural Dementia Care team, and the Rural and Remote Memory Clinic. Thank you to my committee members Drs. Megan O'Connell, Andrew Kirk, and Jamie Campbell as well as to my external examiner Dr. Morris Moscovitch.

To my family and friends, I am forever grateful for everything you have done to help me reach my goals. Thank you for always believing in me.

TABLE OF CONTENTS

LIST OF TABLES

General Introduction

Canada's aging population is placing escalating demands on the health care system (Alzheimer Society of Canada, 2010). As a result, there is an increasing interest in understanding cognitive changes that occur both as a result of healthy, normal aging and as a consequence of dementia. Verbal fluency tests are used frequently in experimental and clinical settings to understand these cognitive changes. Although these neuropsychological measures have been used to assess cognitive performance, comparison of total word production or number of errors (the typical measures used on these tasks) does not reliably differentiate healthy aging and dementia or subtypes of dementia. Further, although executive functioning, semantic memory and metacognition are informative concepts in understanding the cognitive processes that are involved in verbal fluency production, the traditional scoring measure of total words produced does not fully capture these components.

Troyer, Moscovitch, and Winocur (1997) proposed examining two subcomponents of verbal fluency production (i.e., clustering and switching) to further explore the cognitive abilities required for healthy performance on verbal fluency tasks. The current research extended this process approach to interpreting verbal fluency performance by comparing fluency production both across the adult lifespan in groups of healthy individuals and by comparing individuals diagnosed with dementia. Three objectives were identified for this body of research. The first objective was to develop a computerized scoring program to analyze subcomponents of verbal fluency tasks. The second objective of the current research project was to identify the subcomponents of verbal fluency production that are sensitive to age related changes and others that are relatively age stable. The third objective was to determine which subcomponents of verbal fluency production are impacted by dementia and which subcomponents can be used to differentiate dementia subtypes.

Verbal Fluency

Traditional Scoring Measures

Typically, during verbal fluency test administration, participants are given sixty seconds to produce as many words as possible either beginning with a specific letter, on tests of phonemic fluency, or belonging to a given category such as animals, on tests of semantic fluency. The most commonly used scoring measure on these tests is a score of total words produced. In addition to examining total word production on verbal fluency tasks, examination

of the types of errors (i.e., reporting an incorrect exemplar) and perseverations (i.e., repeating a previously generated exemplar) can add useful information about an individual's performance. Three types of perseverations can occur during verbal fluency tests. An individual can repeat previous responses (i.e., recurrent or ideational perseverations), revert back to a previous category (referred to as "stuck in set"), or repeat the same item over and over (i.e., continuous perseveration) (Azuma, 2004). Recurrent perseverations are the most common type of perseveration (Azuma, 2004; Ramage, Bayles, Helm-Estabrooks, & Cruz, 1999). Higher than expected perseveration rates are found in individuals with aphasia, Alzheimer's disease, frontal lobe damage, Parkinson's disease, Huntington's disease, and traumatic brain injury (Azuma, 2004). Intrusions (errors and perseverations) can result from overloading working memory, such as during a dual task condition, especially when the memory load being added is similar to the primary task (Azuma, 2004). This indicates that in addition to effective search and retrieval processes and intact semantic memory stores, verbal fluency performance is also dependent on working memory and and the ability to inhibit intrusion errors.

Cognitive Skills Required For Verbal Fluency Production

Verbal fluency performance is dependent on intact lexical and semantic memory stores for phonemic and semantic fluency, respectively (Gierski, Peretti, & Ergis, 2007). Semantic verbal fluency performance is believed to rely more heavily on temporal lobe functioning (e.g., verbal memory and word storage) whereas phonemic verbal fluency performance is believed to rely more heavily on frontal lobe functions (e.g., strategic search processing and cognitive flexibility; Gierski et al., 2007; Marczinski & Kertesz, 2006; Mummery, Patterson, Hodges, & Wise, 1996). Multiple cognitive components and associated brain regions are needed for normal performance on these tasks, however. For example, both tasks require verbal abilities, search and retrieval skills, adequate speed of processing, and an ability to inhibit inappropriate responses (Abwender, Swan, Bowerman, & Connolly, 2001; Henry & Phillips, 2006).

Memory storage.

Effective semantic verbal fluency performance relies on intact semantic memory, which is associated with medial temporal lobe functioning. Specifically, learning and retrieval aspects of memory are supported by medial temporal lobe systems and associated brain regions, which consolidate memory traces and contribute to the retrieval of information from memory stores (Giovagnoli, Erbetta, Reati, & Bugiani, 2008). As well, the hippocampus is important in

relational learning (i.e., creation and memorization of associations between novel items). Long lasting memories are formed by the hippocampus in an interactive circuit with related limbic structures in the medial temporal lobes and the diencephalon (Moscovitch, 1994). According to this model, these long lasting memories are available when an individual interacts with an appropriate cue at retrieval (Moscovitch, 1994). The ability to store information in semantic memory and retrieve it at a later point depends on proper functioning of the medial temporal lobe structures including the hippocampus. These brain regions appear to be important during semantic verbal fluency production because semantic fluency requires intact conceptual memory and has been shown to rely on intact lateral and inferior temporal lobe regions that are also involved in object perception, recognition, imagery and naming (Rascovsky, Salmon, Hansen, Thal, & Galasko, 2007). Phonemic fluency performance requires retrieving words based on lexical representations, largely mediated by the left prefrontal lobe (Henry & Crawford, 2004; Stuss et al., 1998).

Executive functions.

Verbal fluency performance also depends on components of executive functioning. These components include the ability to search memory for correct words, the ability to shift between words or categories of words, the ability to inhibit inappropriate responses, and working memory (Gierski et al., 2007; Marczinski & Kertesz, 2006; Mummery et al., 1996). The frontal lobes are implicated in working memory, conditional learning, encoding strategies, temporal sequencing, and the retrieval of abstract concepts, all of which contribute to learning and memory (Giovagnoli et al., 2008).

Research into the neurobiological basis of executive functions supports the notion of executive functions constituting distinct but related constructs. The prefrontal cortex is recognized as a critical component of intact executive functioning; however other brain regions are important for input of information (Jurado & Rosselli, 2007). Specifically, the dorsolatoral prefrontal circuit is implicated in planning, goal setting, set-shifting, working memory and self monitoring (Jurado & Rosselli, 2007; Miller, 2007). The lateral orbitofrontal circuit is involved in assessment of risk and inhibition of inappropriate behaviours. (Jurado & Rosselli, 2007) The anterior cingulate circuit functions to monitor behaviour and self-correct errors (Jurado & Rosselli, 2007). Damage to any of these areas is likely to impair performance on verbal fluency tasks. Specifically, phonemic fluency requires the formation of novel categories (e.g., words

starting with a specific letter) and therefore is presumed to require more effort on the part of search processes dependent on intact prefrontal lobe functioning than semantic fluency, where exemplars are already stored categorically in semantic memory (e.g., animal names) (Rascovsky et al., 2007). Both phonemic and semantic fluency place demands on the search and retrieval of information from semantic memory although semantic fluency requires a search of exemplars from a superordinate category (e.g., animals or fruits and vegetables), and thus is dependent on semantic associations, while phonemic fluency requires a less constrained search (e.g., words that start with "F") (Murphy, Rich, & Troyer, 2006). This indicates that both fluency tasks require effective search and retrieval mechanisms, but semantic fluency requires a more complex search from superordinate categories and phonemic fluency requires the effortful formation of novel categories.

Metacognition.

Three metacognitive components are important concepts in understanding verbal fluency production (Young, 2004). Willingness to continue the search, feeling of knowing, and judgement of confidence that the information retrieved from memory is correct, are involved in the process of retrieving answers to questions assessing general knowledge (Young, 2004). These components are likely to be important in retrieving information from semantic memory. Specifically, the willingness to continue searching could become an important factor when an individual begins to have difficulty searching from a specific category (Young, 2004). The feeling of knowing may fluctuate during verbal fluency tasks because, as more words are produced within a specific category, newer items come to mind at a slower rate (Young, 2004). Rewards and penalties are also important in determining an individual's willingness to continue searching. For example, the cost of not recalling items means that an item may not be retrieved; however, the cost of retrieving an incorrect item wastes time and energy (Young, 2004). Young (2004) examined these hypotheses in college students who were asked to generate words from two different natural categories and found that participants spent more time searching categories with high potency (i.e., those categories with a high number of average words generated in thirty seconds). Young (2004) asserted that these results support the theory that participants have a feeling-of-knowing that allows them to judge when to shift categories, and that participants consider the cost and mental effort of switching categories and this evaluation contributes to their willingness to continue searching. This research implies that in addition to executive functioning

and memory storage being integral to the understanding of verbal fluency output, word generation on verbal fluency tasks is also dependent on an individual's evaluation of the costs and benefits of various word generation strategies.

Normal Aging

Cognitive Changes Associated With Age

Verbal fluency performance has been used extensively in previous research to evaluate cognitive changes associated with normal aging. Two prominent views of age related cognitive change dominate the literature: 1) age related cognitive changes are due to a decline in general purpose processing resources, including speed of processing and working memory; and, 2) age related cognitive changes are due to an overall decline in executive functions over and above the effects of processing speed and working memory (Salthouse, 2010). In support of the first hypothesis, with increased age, previous research has shown processing speed declines (Bryan & Luszcz, 2000; Salthouse, 1993; van Hooren et al., 2006). It has been proposed that this age related decline in processing speed is responsible for a large proportion of age related cognitive effects by making it difficult for older adults to rapidly processing information. For example, according to Salthouse (1993; 2010), up to 80% of the variance in age related cognitive change is associated with variations in processing speed. Although increasing age is associated with lower performance on many measures of cognitive functioning, when the effect of declining processing speed is taken into account, the effect of age on these cognitive tasks is markedly decreased (Salthouse, 1993). In addition, Salthouse (1991) found that the removal of the variance accounted for by working memory from tasks examining age related cognitive change further reduces the direct influence of age. Together these results support the contention that many age effects on cognitive tasks are primarily due to age related declines in processing speed and working memory, rather than due to declines in other cognitive functions.

A contrasting view of age related cognitive change asserts that the cognitive functions associated with the prefrontal lobes show age related effects over and above age effects on processing speed and working memory. The prefrontal lobe regions of the brain deteriorate earlier and show larger deficits than other brain regions in older adults (MacPherson, Phillips, & Sala, 2002). The cognitive tasks that have been shown to most consistently evidence decline with increased age are those that require executive functions, which have been linked to intact prefrontal lobe functioning. Specifically, the strategic aspects of encoding and retrieval are

believed to become less efficient with increased age, resulting in age-associated decline in episodic memory (Bryan & Luszcz, 2000). In other words, as individuals age, they have increasing difficulty learning and recalling new information because of declining executive abilities that influence the efficiency of encoding and retrieval processes. As well, planning and organization of behaviour become more difficult with increasing age (van Hooren et al., 2006). Older individuals also show impairment on measures of inhibition which is another component of executive functioning mediated by prefontal lobe structures. On tasks requiring that an individual inhibit specific behaviour, older age groups show poorer performance compared to younger age groups (van Hooren et al., 2006). Overall many components of executive functioning including planning, organization, retrieval and inhibition show age related decline.

Studies examining executive function decline in older adults, however, have yielded variable findings, with some measures of executive functioning not consistently showing age related effects (Jurado & Rosselli, 2007). This variability supports the concept of executive functions as distinct but related components rather than a unitary construct. Age associated changes in executive functions might be limited to specific subtypes of executive functions. For example, within the frontal lobe regions, the dorsolateral prefrontal cortex is implicated in deteriorating executive functions with age. Tasks that are dependent on dorsolateral functioning, such as abstract thinking and problem solving, show more impairment in normal older adults than measures that are dependent on ventromedial prefrontal lobe functioning, such as the regulation of social behaviour (MacPherson et al., 2002). Although this alternative theory for understanding age related cognitive change does not deny that speed of processing and working memory declines with increased age, this theory asserts that there are additional declines in executive functioning that cannot be accounted for solely by examining processing speed and working memory.

Verbal fluency and age.

Age related differences on total word production on verbal fluency tasks commonly are reported. Older adults tend to produce fewer total words on semantic fluency compared to younger age groups (Bryan & Luszcz, 2000; Crossley, D'Arcy, & Rawson, 1997; Haugrud, Lanting, & Crossley, 2010; Lanting, Haugrud, & Crossley, 2009). The relationship between age and performance on phonemic verbal fluency tests appears to be less consistent, with some studies showing age effects (Bryan & Lyszcz, 2000; Crossley et al., 1997; Haugrud et al., 2010;

Lanting et al., 2009) and others not (Brickman et al., 2005; Rodriguez-Aranda & Martinussen, 2006). As well, the effects of education and age on verbal fluency production have been examined. Older individuals with more years of education produce more words on verbal fluency tasks than those with fewer years of education (Kempler, Teng, Dick, Taussig, & Davis, 1998; Mathuranath et al., 2003). Among studies that have found age effects on both semantic and phonemic fluency tests, the pattern of age effects appears to differ between the two types of fluency tasks. For instance, a meta-analytic study of cross sectional studies of phonemic fluency found that phonemic fluency increased until the third decade of life, remained stable during the 40s, then showed a significant decline through the 50s until the late 60s followed by a rapid decline through the 70s and late 80s (Rodriguez-Aranda & Martinussen, 2006). In a cross sectional study, Haugrud and colleagues (2010) found that total word production on phonemic fluency was significantly lower than younger participants by approximately age 60, and remained consistent in the older age groups. In contrast, this same study found semantic fluency production was lowest in an old-old age group (over age 75), followed in order by an old age group (66-74), a middle age group (41-65), and with the highest production in a young age group (20-40). These findings suggest that the pattern of age effects on the two fluency tasks differs and highlights the importance of examining a complete age range rather than comparing young to old individuals.

If the hypothesis that most age related cognitive change is the result of decreased processing speed and working memory (Salthouse, 1993) is correct, we would expect equivalent decline on phonemic and semantic fluency with age since both fluency tasks require intact processing speed and working memory. If the hypothesis that age related cognitive change is the result of decline in executive functioning, over and above processing speed (MacPherson et al., 2002) is correct, we would expect to see a greater age related decline on phonemic fluency compared to semantic fluency because phonemic fluency is relatively more dependent on prefrontal lobe functioning (Gierski et al., 2007; Marczinski & Kertesz, 2006; Mummery et al., 1996). However, in conflict with both of these hypotheses, healthy aging research consistently shows a relatively greater age related decline on semantic fluency compared to phonemic fluency. Results with respect to normal aging and verbal fluency indicate that there could be components of verbal fluency performance that are not assessed in a score of total words produced.

Dementia

In addition to examining healthy cognitive aging, previous research has also examined the relationship between pathological aging (dementia) and verbal fluency test performance. Although this relationship has been examined by numerous previous studies with respect to total word production on verbal fluency tasks, the results have not always been consistent. As well, the reported effects on verbal fluency performance have not always been consistent with the theoretical understanding of the cognitive changes associated with various dementia subtypes.

Alzheimer's Disease (AD)

Approximately 1.5% of the population of Canada (480 600 people) is estimated to suffer from AD, with prevalence expected to increase to 2.8% in the next thirty years (1 125 200 people; Alzheimer Society, 2001). The criteria for the diagnosis of AD emerging from the Third Canadian Consensus Conference on the Diagnosis and Treatment of Dementia (CCCDTD3) include a gradual and progressive onset of declining memory with at least one additional cognitive domain showing impairment (Rockwood, Bouchard, Comiciuli, & Leger, 2007). In addition, cognitive impairments must not be the result of another systemic or neurologic disorder and must be severe enough to cause significant functional impairment (Robillard, 2007). Individuals with AD typically have difficulties with confrontational naming and the most consistently found language deficit in AD is impaired word finding, particularly if given a target semantic category, such as animal names, to guide the speeded generation of words (Braaten, Parsons, McCue, Sellers, & Burns, 2006). Individuals with AD have also been shown to perform relatively lower than other dementia subtypes on measures of memory and learning (Braaten et al., 2006; Giovagnoli et al., 2008).

The disproportionate decline in memory and the learning of new information relative to other cognitive functions in AD results is assumed to result from degeneration of the medial temporal lobe structures including the hippocampus and the adjacent entorhinal cortex (Braaten et al., 2006; Levy & Chelune, 2007; Rascovsky et al., 2007). In addition, AD affects the inferior-lateral temporal lobe resulting in difficulties with spatial processing and accessing semantic knowledge (Hodges et al., 1999). As AD progresses to more advanced stages, the dorsolateral prefrontal cortex becomes impaired and results in impairments in executive functions (Levy & Chelune, 2007).

Verbal fluency and AD.

Examination of the effects of AD on verbal fluency performance supports the medial temporal lobe dysfunction model of AD. Declines in semantic verbal fluency performance have been found in Alzheimer's disease (AD) patients compared to healthy older adults (Crossley et al., 1997; Haugrud et al., 2010; Henry, Crawford, & Phillips, 2004; Mok, Lam, & Chiu, 2004). Phonemic fluency performance also has been shown to decline in AD compared to healthy older adults, but the effect is smaller on the phonemic than on the semantic task (Canning, Leach, Stuss, Ngo, & Black., 2004; Crossley et al., 1997; Haugrud et al., 2010; Henry et al., 2004). Despite a general consistency in the literature, some studies have found that certain subgroups of persons with AD are not more impaired on semantic versus phonemic fluency (Fisher, Tierney, Rourke, & Szalai, 2004). In addition to lower total word generation, individuals with AD produce fewer low frequency exemplars than normal controls (Sailor, Antoine, Diaz, Kuslansky, & Kluger, 2004).

There are two main hypothesized models to explain semantic memory decline in AD. The first model states that semantic knowledge breakdown results from degradation in the structure or content of semantic memory (Henry et al., 2004; Sailor et al., 2004). In contrast, the second model proposes that deficits in semantic memory in AD reflect a deficit in the cognitive processes that are responsible for accessing semantic knowledge (i.e., the executive control mechanisms responsible for memory retrieval), while the semantic store itself remains intact (Henry et al., 2004). Henry et al. concluded that studies of verbal fluency support the semantic storage breakdown hypotheses. Semantic fluency is more dependent than phonemic fluency on an intact semantic store, as shown by the higher correlation between semantic fluency and measures of semantic storage such as the Boston Naming Test (Henry et al., 2004). Phonemic fluency, in contrast, is not correlated with measures of semantic storage. As well, both phonemic and semantic fluency have been shown to be related to intact executive functioning (Henry et al., 2004). Since individuals with AD typically have a larger decline on semantic fluency than phonemic fluency, this supports the hypothesis that semantic memory decline is more prominent than executive function decline in early stage AD. This model implies that individuals with AD have a smaller set of items to search to generate words on fluency tasks and consequently produce fewer words.

Mild Cognitive Impairment (MCI)

The Canadian Study of Health and Aging reported a 16.8% prevalence of mild cognitive impairment not meeting criteria for dementia (CIND) in those over age 65 (Chertkow et al., 2007). As well, it has been estimated that as high as 44% of individuals diagnosed with mild cognitive impairment (MCI) convert to a diagnosis of AD after three years (Chertkow et al., 2007), making a diagnosis of MCI important in understanding preclinical cognitive decline in older adults. Amnestic mild cognitive impairment (aMCI) is characterized by selective memory impairment with the preservation of functional abilities in daily life (Chertkow et al., 2007; Lam, Lui, Chiu, Chan, & Tam, 2005; Petersen et al., 1999). These individuals have been shown to have lower scores on measures of delayed recall, digits backwards, and visual span compared to controls (Lam et al., 2005), but this impairment is less severe than in individuals diagnosed with AD.

Verbal fluency and MCI.

Individuals with MCI have also been found to demonstrate declines on semantic relative to phonemic fluency even though their overall performance was within normal range (Murphy et al., 2006). Some studies report impaired semantic (Nutter-Upham et al., 2008; Raoux et al., 2008; Fagundo et al., 2008) and phonemic (Nutter-Upham et al., 2008) total word production in MCI groups compared to healthy older adults, while other studies have failed to show phonemic or semantic total word decline in MCI (Murphy et al., 2006).

Vascular Dementia (VaD)

The term Vascular Cognitive Impairment (VCI) is used increasingly in clinical and research settings as a broader term encompassing all forms of cognitive loss due to cerebrovascular disease (Rockwood et al., 2007). VCI-no dementia, subcortical vascular dementia (VaD) with white matter changes on neuroimaging, and VaD with multiple or single infarcts are three recognized subtypes of VCI (Rockwood et al., 2007). The National Institute of Neurologic Disorders and Stroke and the Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN) criteria for VaD require a diagnosis of dementia (including decline from a previous level of functioning and impairment on memory and two or more other cognitive domains), evidence of cerebrovascular disease, and a relationship between dementia presentation and cerebrovascular disease (Roman et al., 1993). However, the CCCDTD3 reported that these criteria show high specificity at a cost of low sensitivity

(Robillard, 2007). Since VaD can occur as the result of a single brain infarct or multiple discrete smaller lesions (Robillard, 2007), the anatomical profile of VaD is variable depending on the nature of the associated cerebrovascular disease. For example, this dementia can result from primarily cortical, subcortical or a combination of cortical and subcortical damage (Braaten et al., 2006), although the first structural changes in VaD are typically seen in the fronto-striatal circuitry (Jones, Laukka, & Backman, 2006). Given the multiple potential neuroanatomical causes of VCI, a consistent neuropsychological profile for this disease is not expected, although there tend to be some general commonalities across individuals diagnosed with VCI. For example, previous research has supported a dysexecutive profile of VaD, including a general slowing in cognitive performance (Lafosse et al., 1997; Levy & Chelune, 2007; Robillard, 2007; Rockwood et al., 2007). When compared to individuals with frontotemporal dementia, individuals with VaD tend to perform worse on measures of memory, although those with VaD show equivalent decline on memory measures when compared to individuals with AD (Braaten et al., 2006). Importantly, this decline in memory in VaD is attributed to deficits in retrieval from the semantic store rather than from decay in the store, which is presumed to occur in individuals with AD.

Verbal fluency and VaD.

Individuals with vascular dementia show significantly lower word production on both phonemic and semantic fluency tasks compared to healthy older adults (Braaten et al., 2006). In contrast to individuals with AD, individuals with VaD show relatively equivalent deficits on both fluency tasks. As a result of this equivalent decline on both tasks in VaD, individuals with VaD have lower output than individuals with AD on phonemic fluency tasks (Canning et al., 2004; Lafosse et al., 1997; Levy & Chelune, 2007). Individuals diagnosed with vascular cognitive impairment – no dementia, often considered a precursor to VaD, have been shown to have slight but nonsignificant deficits on phonemic fluency (but not in semantic fluency) when compared to controls (Canning et al., 2004). Overall, research generally shows that both phonemic and semantic fluency total word production show impairment in VaD, and that production on phonemic fluency appears to be impacted at the earliest stages.

Frontotemporal Dementia (FTD)

The average age of onset of FTD is the late 50s and this dementia accounts for approximately 10-20% dementias with a higher percentage of early dementias (Wittenberg et al.,

2008). The Neary criteria (Neary et al., 1998) are most commonly used in diagnosing FTD and include onset of symptoms before age 65, insidious onset with gradual progression, early decline in interpersonal conduct, early difficulties regulating personal conduct, emotional blunting and loss of insight. Supportive features include loss of personal hygiene, mental rigidity, hyperorality, and perseverative behaviour. Individuals with FTD tend to have greater functional impairment than individuals with AD even when individuals with FTD perform at equivalent or higher levels on cognitive screening measures than those with AD (Wittenberg et al., 2008). FTD results from degeneration of the prefrontal and anterior temporal lobes which are responsible for reasoning, personality, speech, language and some parts of memory (Braaten et al., 2006). Marked changes in personality and behaviour including apathy, irritability, disinhibition, poor insight and lack of social awareness reflect early degradation of the orbitofrontal cortex and the network involving the insula, striatum and medial frontal lobes (Wittenberg et al., 2008).

FTD is characterized by rigid and inflexible thinking, impaired judgement, and impaired executive functions with relatively preserved memory. For example, episodic memory tends to be well preserved in FTD, while measures of executive functioning (e.g., perseveration, rule violations) tend to be more impaired than in individuals with AD (Wittenberg et al., 2008). These individuals show impaired performance on the Trail-Making test, a measure of mental set shifting and processing speed, as well as difficulties on measures of attention (Braaten et al., 2006). Individuals with FTD have difficulties organizing strategies to encode information, which results in impairments on measures of free recall and recognition. On measures of delayed recall, visuoconstruction and word list learning, individuals with FTD perform better than those with AD (Diehl & Kurz, 2002).

FTD can be divided into three subtypes. Frontotemporal dementia – behavioural variant (FTD-bv) is characterized by a marked disturbance in personality and social conduct, which reflects the orbitobasal frontal lobe focus of degeneration (Robillard, 2007). Progressive nonfluent aphasia (FTD-pnf) is characterized by progressive decline in fluent speech resulting in halting speech with lexical, phonological and syntactic deficits, although comprehension of language and repetition remain relatively preserved (Wittenberg et al., 2008). FTD-pnf results from left inferior frontal and insular atrophy, particularly around the perisylvian cortex (Giovagnoli et al., 2008; Levy & Chelune, 2007; Wittenberg et al., 2008). In semantic dementia (FTD-SD) individuals lose the semantic meaning of words which results in anomia, impaired

comprehension and fluent but empty spontaneous speech (Davies et al., 2005; Hodges et al., 1999; Wittenberg et al., 2008). FTD-SD results from anterior temporal lobe damage, specifically the anterior parahippocampal and fusiform regions including the perirhinal cortex with typically more severe damage to the left hemisphere (Davies et al., 2005). Notably, individuals with SD show maintained episodic memory and visuospatial skills (Marczinski & Kertesz, 2006).

Verbal fluency and FTD.

Verbal fluency performance has been investigated both by combining all cases of FTD and by examining individual subtypes of FTD. Decreased word production on both semantic and phonemic fluency tasks has been found in individuals with frontotemporal dementia, with more severe impairment on the phonemic task (Rascovsky et al., 2007). Impaired verbal fluency performance in FTD has been associated with general adynamia (i.e., loss of strength) and deficits in motor responses (Diehl & Kurtz, 2002) as well as deficits in retrieval processes (Rascovsky et al., 2007). Individuals with FTD-SD and FTD-pnf show more impairment on measures of verbal fluency than FTD-bv or AD individuals, particularly on phonemic fluency measures (Levy & Chelune, 2007). Individuals with FTD-pnf tend to produce the fewest words on verbal fluency tasks followed by FTD-SD and individuals with AD (Marczinski & Kertesz, 2006). As well, the words produced on fluency tasks by individuals with SD tend to be high frequency words as compared to controls and those with AD (Marczinski & Kertesz, 2006). Overall, all subtypes of FTD tend to show impairment on verbal fluency measures with a larger impairment being evident on the phonemic test, likely due to the higher retrieval demands of this task.

Dementia with Lewy Bodies (DLB)

Dementia with Lewy Bodies accounts for 10-25% of all dementia cases or approximately 2% of individuals over age 65 (Troster, 2008; Oda, Yamamoto, & Maeda, 2009). The core features of DLB are fluctuations in attention and alertness, visual hallucinations that are well formed and detailed, and spontaneous motor features of Parkinsonism (Robillard, 2007). Visual hallucinations, delusions, auditory hallucinations and olfactory hallucinations occur in approximately 54%, 49%, 25%, and 7% of cases respectively (Levy & Chelune, 2007). Cognitive abilities can fluctuate markedly in DLB due to fluctuations in attention and alertness. DLB is characterized by deficits in visualspatial ability, attention, speed of processing and executive functioning (Levy & Chelune, 2007; Oda et al., 2009). Visual spatial impairments and

attentional deficits are larger in DLB than in AD (Oda et al., 2009; Ralph, Howard, Whitworth, Garrard, & Hodges, 2001).Memory remains relatively preserved in the early stages of this disorder although individuals with DLB show poor initial acquisition of information. Individuals with DLB perform below individuals with AD on measures of perception, planning and organization, attention, phonemic fluency and divided attention (Levy & Chelune, 2007).

Verbal fluency and DLB.

Persons with DLB and AD tend to show equivalent impairment on measures of semantic fluency, although individuals diagnosed with DLB show more impairment than those with AD on phonemic fluency (Levy & Chelune, 2007). This difference on fluency measures in DLB is hypothesized to result from the greater demand phonemic fluency places on executive functions. Individuals with DLB tend to show equal levels of impairment on phonemic and semantic fluency, likely resulting from poor executive and working memory functions (Ralph et al., 2001).

The Two Component Model of Verbal Fluency

Troyer and Colleagues (1997) Model

Verbal fluency tests have been examined extensively for their utility in understanding healthy aging and dementia through a score of total words produced. A process approach to neuropsychological test interpretation is an alternative approach to assessment beyond simply comparing total scores on measures (Milberg, Hebben, & Kaplan, 2009). This approach examines the cognitive components of a task required for normal performance. Through this method specific strategies and approaches to a task can be compared to provide additional information over and above group differences on total scores. Troyer and colleagues (1997) proposed a two component model of verbal fluency production which is an example of this approach. Specifically, Troyer and colleagues (1997) proposed that verbal fluency performance can be divided into two components: 1) clustering, or the production of words within a semantic or phonemic subcategory; and 2) switching, or the ability to shift between clusters. According to these authors, verbal fluency performance depends on the search for appropriate subcategories and the production of words within these categories. Clustering is proposed to rely on temporal lobe processes to produce exemplars of a category and switching is proposed to rely on frontal lobe functions for strategic search processes. Both clustering and switching have been shown to be highly correlated with semantic total word production while phonemic production is correlated only with switching (Troyer et al., 1997).

Support for this model has been found in brain lesion studies, with individuals with frontal lobe lesions showing impaired switching rates and individuals with temporal lobe lesions showing diminished semantic cluster size (Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998). Recently, brain imaging studies have been conducted to examine clustering and switching. Hirshorn and Thompson-Schill (2006) found activation of the left inferior frontal gyrus using fMRI when participants switched between subcategories on semantic verbal fluency. In addition, in this study bilateral temporal regions showed greater activation between switches, the portion of the task when an individual would be relying on clustering strategies (Hirshorn & Thompson-Schill, 2006). These studies support the two component model of performance during verbal fluency tasks, including the role of frontal lobe functions in switching and the role of temporal lobe functions in clustering.

Age effects.

Older adults switch less frequently on verbal fluency tasks than younger groups (Bruicki & Rocka, 2004; Haugrud et al., 2010; Lanting et al., 2009; Troyer et al., 1997; 2000). Since aging is believed to be associated with decreases in executive functioning, this is consistent with the two component model of verbal fluency (Henry & Phillips, 2006). In contrast, some studies have reported that older adults produce larger phonemic clusters than younger groups (Troyer et al., 1997; 2000; Hughes & Bryan, 2002). However, studies that have found a clustering advantage in older adults tend to describe older adults with atypically high education levels, which may artificially advantage older adults on verbal fluency by creating cohort differences in general verbal ability (Crossley et al., 1997). As a result of these sampling differences, reports of the effect of age on clustering and switching have been inconsistent.

Dementia effects.

Clustering and switching strategies also have been examined in individuals diagnosed with Alzheimer's disease. According to the two component model of Troyer and colleagues (1997), individuals with AD should show smaller cluster sizes with relatively intact switching rates, due to AD-related decreases in semantic knowledge. These results have been found by some researchers (Haugrud et al., 2010; Troyer, Moscovitch, Winocur, Leach, & Freedman, 1998). However, other studies have found that patients with Alzheimer's disease tend to produce both fewer switches and smaller cluster sizes on verbal fluency tests (Troster et al., 1998; Beatty, Testa, English, & Winn, 1997; Fagundo et al., 2008; Gomez & White, 2006). Methodological

inconsistencies among studies may account for these inconsistent results. In studies that have used clustering and switching measures in groups of participants with Alzheimer's disease, the stage of the disease of participants varies across studies, with some studies recruiting individuals in more advanced stages (Epker, Lacritz, & Munro Cullum, 1999; Troster et al., 1998) and other studies using mild or early stage AD patients (Beatty et al., 1997; Haugrud et al., 2010; Troyer, Moscovitch, Winocur, Leach, et al., 1998). This could result in differences across studies because individuals at a more severe stage of AD produce fewer total words, which can limit the measurement of clustering and switching components. In addition, as the disease progresses, declines in other cognitive processes will become more pronounced, influencing both clustering and switching scores. Alternatively it is possible that the scores used to measure clustering and switching as proposed by Troyer and colleagues (1997) do not fully capture the cognitive abilities required for verbal fluency production.

Abwender and Colleagues (2001) Model

Abwender and colleagues (2001) proposed modifications to the scoring procedures used to assess the two component model of verbal fluency, and specifically proposed assessing two types of switches. Hard switches occur between two single, non-clustered words or between a clustered word and a single word and are believed to reflect the speeded nature of verbal fluency tasks. Cluster switching occurs between two groups of clustered words and is believed to reflect mental flexibility.

Additional Modifications to the Two Component Model

March and Pattison (2006) proposed examining the raw number of subcategories used by individuals during semantic verbal fluency performance. This variable was proposed to examine the access of individuals to multiple subcategories during word generation. In that regard, they found that individuals with AD access fewer subcategories than healthy controls on semantic fluency tasks (March & Pattison, 2006).

Mayr (2002) proposed that an individual's score on number of switches in the Troyer and colleagues (1997) model can be impaired either because the individual has difficulties accessing new semantic clusters or they have difficulty generating words within clusters. Further, this author proposed that switching deficits might indicate generally slowed retrieval, both within or between clusters. In addition, Ross and colleagues (2007) argued that clustering on phonemic fluency tasks may be an artefact of the test itself rather than a deliberate, strategic process. To

summarize Ross and colleagues (1997), it may be that the speeded nature of verbal fluency tasks results in clustering and switching, rather than clustering and switching being overt strategies as proposed by the model of Troyer and colleagues (1997).

To address some of the concerns of previous researchers, Lanting and colleagues (2009) proposed additional scoring procedures to the methods of Troyer et al. (1997) and Abwender et al. (2001). These authors examined the number of novel and repeated clusters produced by healthy young and older participants and proposed that executive dysfunction associated with older age would produce both fewer novel clusters, due to deficits in the retrieval process, and more repeated clusters, due to increased difficulty with inhibition (i.e., repeating previously used clusters was proposed to be a less advantageous strategy). Lanting and colleagues (2009) found that younger adults produced both more novel clusters and more repeated clusters, which the authors interpreted as indicating that returning to repeated clusters may actually be a beneficial strategy. As well, these authors examined the percentage of clustered words. This variable was included to address limitations of the Troyer and colleagues (1997) model that included single words as a cluster with a score of zero. In the study by Lanting and colleagues (2009) however, percentage of clustered words failed to differentiate between younger and older age groups, suggesting that this variable might be more applicable in differentiating individuals with dementia from healthy older adults due to expected differences in cluster size as opposed to showing age related effects.

Haugrud and colleagues (2010) also attempted to address limitations of the Troyer and colleagues (1997) method of calculating clustering and switching scores. In the original study by Troyer and colleagues (1997), perseverations and errors were included because the researchers proposed that these intrusions might be assisting with the strategy use of individuals on verbal fluency tasks, prompting them to initiate new clusters. However, evidence was not provided by Troyer et al. for this assertion that perseverations are strategic rather than a random occurrence throughout word production. If perseverations and errors do not occur systematically, as Troyer and colleagues (1997) propose, but rather are randomly distributed, then including these intrusions might bias the assessment of verbal output in clustering and switching. Particularly, this may artificially alter the cluster size and switching scores for older individuals and individuals with Alzheimer's disease who tend to produce more errors and perseverations than healthy younger individuals. Haugrud and colleagues (2010) examined total word production,

and clustering and switching on verbal fluency tasks, both with errors and perseverations included and with these intrusions excluded. These researchers found that errors and perseverations were not systematic and that their inclusion in the calculation of clustering and switching scores did artificially inflate the cluster size scores of individuals with AD. When these intrusions were removed, individuals diagnosed with AD produced smaller cluster sizes on verbal fluency tasks compared to healthy older adults, which is consistent with the two component model of verbal fluency.

Limitations of Previous Research on the Two Component Model

Although clustering and switching have been investigated in groups with Alzheimer's disease, these subcomponents of verbal fluency output have received limited attention in other dementia subtypes. Currently no studies examining vascular dementia, Dementia with Lewy Bodies, or frontotemporal dementia have been conducted using measures of clustering and switching. Since total word production on verbal fluency tasks has been found to inconsistently differentiate these diagnostic groups, examination of clustering and switching strategies can potentially offer a more consistent method of differentiating dementia subgroups.

Overview of Current Research

The objective of the current body of research was to investigate verbal fluency performance both in healthy aging and in subtypes of dementia. The goal of study 1 was to develop a computer program to score clustering and switching measures on verbal fluency tests. It was hypothesized that the computerized scoring program would reliably score clustering and switching measures compared to hand scoring and that computerized scoring would reduce the time required for scoring.

Study 2 compared clustering and switching variables in a healthy aging sample. Individuals were compared across three age groups (young adults, 20-38 yrs; middle-aged adults, 40-63 yrs, and older adults, 65-82 yrs) on measures of clustering and switching, as described by Troyer and colleagues (1997), Abwender and colleagues (2001), and Lanting and colleagues (2009). Based on previous research and the hypothesis of executive functioning decline in healthy aging, it was hypothesized that for both fluency tasks, when compared to the young and middle-age groups, the oldest age group would produce fewer total words and switches. No age category effects were hypothesized for average cluster size or for percentage of clustered words because memory storage is presumably intact in healthy aging.

Study 3 compared verbal fluency performance (using the variables described by Abwender et al., 2001, Lanting et al., 2001, and Troyer et al., 1997) in a group diagnosed with probable Alzheimer's disease to a group of healthy older adults. Based on previous research and hypothesized decline in medial temporal lobe integrity in early stage Alzheimer's disease, it was hypothesized that the AD group would produce fewer total words and smaller average cluster size scores on both verbal fluency tasks when compared to the healthy older adult group, with a larger effect observed on the semantic fluency tasks compared to the phonemic task. Hard switching was predicted to remain intact in the early stage AD group but the AD group was predicted to produce fewer cluster switches, resulting in reduced total switching compared to healthy older adults.

Study 4 compared verbal fluency performance longitudinally over two or three years in individuals diagnosed with probable Alzheimer's disease. Measures of clustering and switching as described by Abwender and colleagues (2001), Lanting and colleagues (2009) and Troyer and colleagues (1997) were used to compare performance across time. It was hypothesized that individuals diagnosed with probable AD would show decline on phonemic and semantic fluency total words over the one and two year follow up periods, with a larger decline on semantic fluency, consistent with previous research. As well, participants were hypothesized to show a decline in semantic fluency average cluster size over time but no change in phonemic average cluster size due to increased disease effects on the medial temporal lobe. With disease progression, individuals diagnosed with Alzheimer's disease were predicted to show longitudinal decline on variables hypothesized to depend more on prefrontal lobe structures (i.e. switching and novel cluster access).

The goal of study 5 was to compare verbal fluency performance across subtypes of dementia. Clustering and switching measures (Abwender et al., 2001; Lanting et al., 2009; Troyer et al., 1997) were compared in groups diagnosed with amnestic mild cognitive impairment (aMCI), Alzheimer's disease (AD), Vascular dementia (VaD), Dementia with Lewy Bodies (DLB), behavioural variant frontotemporal dementia (FTD-bv), and language variant frontotemporal dementia (FTD-lang). Performance was compared across groups and to a healthy older adult group. It was hypothesized the aMCI and AD groups, compared to normal age-equivalent adults would show impaired performance on measures sensitive to medial temporal lobe integrity (i.e. semantic fluency total words, average cluster size). The FTD-bv group was

expected to show impaired performance on measures sensitive to prefrontal lobe functioning (i.e. phonemic fluency total word production, total switches). The FTD-lang group was hypothesized to show the largest fluency decline compared to a healthy control group on all measures. The VaD and DLB groups were hypothesized to show equivalent decline on the phonemic and semantic tasks and impaired switching, but intact cluster sizes, due to hypothesized subcortical impairment.

Together this body of research will provide important insight into the aspects of verbal fluency production that are sensitive to age effects and which aspects remain age stable. As well, this research will show which verbal fluency components are sensitive to subtypes of dementia and how fluency production changes over time in Alzheimer's disease.

References

Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment, 8*, 323-336.

Alzheimer Society (2010). *Rising tide: The impact of dementia on Canadian society*. Retrieved from Alzheimer Society website: http://www.alzheimer.ca/english/rising_tide/rising_tide.htm

Azuma, T. (2004). Working memory and perseveration in verbal fluency. *Neuropsychology, 18*, 69-77.

Beatty, W. W., Testa, J. A., English, S., & Winn, P. (1997). Influences of clustering and switching on the verbal fluency performance of patients with Alzheimer's disease. *Aging, Neuropsychology, and Cognition, 4*, 273-279.

Braaten, A. J., Parsons, T. D., McCue, R., Sellers, A., & Burns, W. J. (2006). Neurocognitive differential diagnosis of dementing diseases: Alzheimer's dementia, vascular dementia, frontotemporal dementia, and major depressive disorder. *International Journal of Neuroscience, 116*, 1271-1293. doi: 10.1080/00207450600920928

Brickman, A. M., Paul, R. H., Cohen, R. A., Williams, L. M., MacGregor, K. L., Jefferson, A. L., … Gordon, E. (2005). Category and letter verbal fluency across the adult lifespan: Relationship to EEG theta power. *Archives of Clinical Neuropsychology, 20*, 561-573. doi: 10.1016/j.acn.2004.12.006

Bryan, J., & Luszcz, M. A. (2000). Measurement of executive function: Consideration for detecting adult age differences. *Journal of Clinical and Experimental Neuropsychology, 22*, 40-55. doi: 10.1076/1380-3395%28200002%2922:1;1-8;FT040

Canning, S. J., Leach, L., Stuss, D., Ngo, L., & Black, S. E. (2004). Diagnostic utility of abbreviated fluency measures in Alzheimer disease and vascular dementia. Neurology, 62, 556-562.

Chertkow, H., Nasreddine, Z., Joanette, Y., Drolet, V., Kirk, J., Massoud, F., Belleville, S., & Bergman, H. (2007). Mild cognitive impairment and cognitive impairment, no dementia: Part A, concept and diagnosis. *Alzheimer's & Dementia, 3*, 266-282.

Crossley, M., D'Arcy, C., & Rawson, N. (1997). Letter and category fluency in community-dwelling Canadian seniors: A comparison of normal participants to those with dementia

of the Alzheimer or vascular type. *Journal of Clinical and Experimental Neuropsychology, 19*, 52-62. doi: 10.1080/01688639708403836

Davies, R. R., Hodges, J. R., Kril, J. J., Patterson, K., Halliday, G. M., & Xuereb, J. H. (2005). The pathological basis of semantic dementia. *Brain, 128*, 1984-1995. doi: 10.1093/brain/awh582

Diehl, J., & Kurz, A. (2002). Frontotemporal dementia: Patient characteristics, cognition, and behaviour. *International Journal of Geriatric Psychiatry, 17*, 914-918. doi: 10.1002/gps.709

Epker, M. O., Lacritz, L. H., & Munro Cullum, C. (1999). Comparative analysis of qualitative verbal fluency performance in normal elderly and demented populations. *Journal of Clinical and Experimental Neuropsychology, 21*, 425-434.

Fagundo, A. B., Lopez, S., Romero, M., Guarch, J., Marcos, T., & Salamero, M. (2008). Clustering and switching in semantic fluency: Predictors of the development of alzheimer's disease. *International Journal of Geriatric Psychiatry, 23*, 1007-1013. doi: 10.1002/gps.2025

Fisher, N. J., Tierney, M. C., Rourke, B. P., & Szalai, J. P. (2004). Verbal fluency patterns in two subgroups of patients with Alzheimer's disease. *The Clinical Neuropsychologist, 18*, 122-131.

Gierski, F., Peretti, C. S., & Ergis, A. M. (2007). Effects of the dopamine agonist piribedil on prefrontal temporal cortical network function in normal aging as assessed by verbal fluency. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 31*, 262-268. doi: 10.1016/j.pnpbp.2006.06.017

Giovagnoli, A. R., Erbetta, A., Reati, F., & Bugiani, O. (2008). Differential neuropsychological patterns of frontal variant frontotemporal dementia and Alzheimer's disease in a study of diagnostic concordance. *Neuropsychologia, 46*, 1495-1504.

Gomez, R. G., & White, D. A. (2006). Using verbal fluency to detect very mild dementia of the Alzheimer type. *Archives of Clinical Neuropsychology, 21*, 771-775. doi: 10.1016/j.acn.2006.06.012

Haugrud, N., Lanting, S., & Crossley, M. (2010). The effects of age, sex and Alzheimer's disease on strategy use during verbal fluency tasks. *Aging, Neuropsychology, & Cognition, 17*. 220-239. doi: 10.1080/13825580903042700

Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology, 18*, 284-295. doi: 10.1037/0894-4105.18.2.284

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia, 42*, 1212-1222. doi: 10.1016/j.neuropsychologia.2004.02.001

Henry, J. D., & Phillips, L. H. (2006). Covariates of production and perseveration on tests of phonemic, semantic and alternating fluency in normal aging. *Aging, Neuropsychology, and Cognition, 13*, 529-551. doi: 10.1080/138255890969537

Hirshorn, E. A., & Thompson-Schill, S. L. (2006). Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency. *Neuropsychologia, 44*, 2547-2557. doi: 10.1016/j.neuropsychologia.2006.03.035

Hodges, J. R., Patterson, K., Ward, R., Garrard, P., Bak, T., Perry, R., … Gregory, C. (1999). The differentiation of semantic dementia and frontal lobe dementia (temporal and frontal variants of frontotemporal dementia) from early Alzheimer's disease: A comparative neuropsychological study. *Neuropsychology, 13*, 31-40. doi: 10.1037/0894-4105.13.1.31

Hughes, D. L., & Bryan, J. (2002). Adult age differences in strategy use during verbal fluency performance. *Journal of Clinical and Experimental Neuropsychology, 24*, 642-654. doi: 10.1076/jcen.24.5.642.1002

Jones, S., Laukka, E. J., & Backman, L. (2006). Differential verbal fluency deficits in the preclinical stages of Alzheimer's disease and vascular dementia. *Cortex, 42*, 347-355.

Juardo, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review, 17*, 213-233. doi: 10.1007/s11065-007-9040-z

Kempler, D., Teng, E. L., Dick, M., Taussig, M., & Davis, D. (1998). The effects of age, education, and ethnicity on verbal fluency. *Journal of the International Neuropsychological Society, 4*, 531-538.

Lafosse, J., Reed, B., Mungas, D., Sterling, S., Wahbeh, H., & Jagust, W. (1997). Fluency and memory differences between ischemic vascular dementia and Alzheimer's disease. *Neuropsychology, 11*, 514-522. doi: 10.1037/0894-4105.11.4.514

Lam, L. C. W., Lui, V. W. C., Chiu, H. F. K., Chan, S. S. M., & Tam, C. W. C. (2004). Executive function impairment in community elderly subjects with questionable dementia. *Dementia and Geriatric Cognitive Disorders, 19*, 86-90. doi: 10.1159/000082354

Lanting, S., Haugrud, N., & Crossley, M. (2009). The effects of age and sex on clustering and switching during speeded verbal fluency tasks. *Journal of the International Neuropsychological Society, 15*, 196-204. doi: 10.1017/S1355617709090237

Levy, J. A., & Chelune, G. J. (2007). Cognitive-behavioral profiles of neurodegenerative dementias: Beyond Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology, 20*, 227-236. doi: 10.1177/0891988707309906

MacPherson, S.E., Phillips, L. H., & Della Sala, S. (2002). Age, executive function, and social decision making: A dorsolateral prefrontal theory of cognitive aging. *Psychology and Aging, 17*, 598-609. doi: 10.1037/0882-7974.17.4.598

March, E. G., & Pattison, P. (2006). Semantic verbal fluency in Alzheimer's disease: Approaches beyond the traditional scoring system. *Journal of Clinical and Experimental Neuropsychology, 28*, 549-566. doi: 10.1080/13803390590949502

Marczinski, C. A., & Kertesz, A. (2006). Category and letter fluency in semantic dementia, primary progressive aphasia, and Alzheimer's disease. *Brain and Language, 97*, 258-265. doi: 10.1016/j.bandl.2005.11.001

Mathuranath, P. S., George, A., Cherian, P. J., Alexander, A., Sarma, S. G., & Sarma, P. S. (2003). Effects of age, education and gender on verbal fluency. *Journal of Clinical and Experimental Neuropsychology, 25*, 1057-1064.

Mayr, U. (2002). On the dissociation between clustering and switching in verbal fluency: Comment on Troyer, Moscovitch, Winocur, Alexander, and Stuss. *Neuropsychologia, 40*, 562-566.

Milberg, W. P., Hebben, N., & Kaplan, E. (2009). The Boston Process Approach to neuropsychological assessment. In I. Grant, & K. Adams (Eds.), *Neuropsychological Assessment of Neuropsychiatric Disorders, 3rd Edition*. New York: Oxford University Press.

Miller, B. L. (2007). The human frontal lobes. In B. L. Miller, & J. L. Cummings (Eds.), *The Human Frontal Lobes: Functions and Disorders* (pp. 3-11). New York: The Guilford Press.

Mok, E. H. L., Lam, L. C. W., & Chiu, H. F. K. (2004). Category verbal fluency test performance in Chinese elderly with Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 18,* 120-124. doi: 10.1159/000079190

Moscovitch, M. (1994). Cognitive resources and dual-task interference effects at retrieval in normal people: The role of the frontal lobes and medial temporal cortex. *Neuropsychology, 8*, 524-534.

Mummery, C., Patterson, K., Hodges, J., & Wise, R. (1996). Generating "tiger" as an animal name or a word beginning with T: Differences in brain activation. *Proceedings: Biological Sciences, 263*, 989-995.

Murphy, K. J., Rich, J. B., & Troyer, A. K. (2006). Verbal fluency patterns in amnestic mild cognitive impairment are characteristic of Alzheimer's type dementia. *Journal of the International Neuropsychological Society, 12*, 570-574. doi: 10.1017/S1355617706060590

Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., … Benson, D. F. (1998). Frontotemporal lobar degeneration: A consensus on clinical diagnostic criteria. *Neurology, 51*, 1546-1554.

Nutter-Upham, K. E., Saykin, A. J., Rabin, L. A., Roth, R. M., Wishart, H. A., Pare, N., & Flashman, L. A. (2008). Verbal fluency performance in amnestic MCI and older adults with cognitive complaints. *Archives of Clinical Neuropsychology, 23*, 229-241. doi: 10.1016/j.acn.2008.01.005

Oda, H., Yamamoto, Y., & Maeda, K. (2009). The neuropsychological profile in dementia with Lewy bodies and Alzheimer's disease. *International Journal of Geriatric Psychiatry, 24*, 125-131. doi: 10.1002/gps.2078

Ralph, M. A., Howard, D., Whitworth, A. B., Garrard, P., & Hodges, J. R. (2001). Semantic memory is impaired in both dementia with Lewy bodies and dementia of the Alzheimer's type: A comparative neuropsychological study and literature review. *Journal of Neurology, Neurosurger, & Psychiatry, 70*, 149-156. doi: 10.1136/jnnp.70.2.149

Ramage, A., Bayles, K., Helm-Estabrooks, N., & Cruz, R. (1999). Frequency of perseveration in normal subjects. *Brain and Language, 66,* 329-340.

Raoux, N., Amieva, H., Le Goff, M., Auriacombe, S., Carcaillon, L., Letenneur, L., et al. (2008). Clustering and switching processes in semantic verbal fluency in the course of Alzheimer's disease subjects: Results from the PAQUID longitudinal study. *Cortex, 44,* 1188-1196. doi: 10.1016/j.cortex.2007.08.019

Rascovsky, K., Salmon, D. P., Hansen, L. A., Thal, L. J., & Galasko, D. (2007). Disparate letter and semantic category fluency deficits in autopsy-confirmed frontotemporal dementia and Alzheimer's disease. *Neuropsychology, 21,* 20-30. doi: 10.1037/0894-4105.21.1.20

Robillard, A. (2007). Clinical diagnosis of dementia. *Alzheimer's & Dementia, 3*, 292-298. doi: 10.1016/j.jalz.2007.08.002

Rockwood, K., Bouchard, R. W., Camicioli, R., & Leger, G. (2007). Toward a revision of criteria for the dementias. *Alzheimer's & Dementia, 3*, 428-440. doi: 10.1016/j.jalz.2007.07.014

Rodriguez-Aranda, C., & Martinussen, M. (2006). Age-related differences in performance of phonemic verbal fluency measured by Controlled Oral Word Association Task (COWAT): A meta-analytic study. *Developmental Neuropsychology, 30,* 697-717. doi: 10.1207/s15326942dn3002_3

Ross, T., Calhoun, E., Cox, T., Wenner, C., Kono, W., & Pleasant, M. (2007). The reliability and validity of qualitative scores for the Controlled Oral Word Association Test. *Archives of Clinical Neuropsychology, 22*, 475-488. doi: 10.1016/j.acn.2007.01.026

Sailor, K., Antoine, M., Diaz, M., Kuslansky, G., & Kluger, A. (2004). The effects of Alzheimer's disease on item output in verbal fluency tasks. *Neuropsychology, 18,* 306-314. doi: 10.1037/0894-4105.18.2.306

Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science, 2*, 179-183.

Salthouse, T. A. (1993). Speed mediation of adult age differences in cognition. *Developmental Psychology, 29,* 722-738.

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society, 16,* 754-760. doi: 10.1017/S1355617710000706

Stuss, D. T., Alexander, M. P., Hamer, L., Palumbo, C. Dempster, R., Binns, M., … Izukawa, D. (1998). The effects of focal anterior and posterior brain lesions on verbal fluency. *Journal of the International Neuropsychological Society, 4*, 265-278.

Troster, A. I. (2008). Neuropsychological characteristics of Dementia with Lewy Bodies and Parkinson's disease with dementia: Differentiation, early detection, and implications for "mild cognitive impairment" and biomarkers. *Neuropsychology Review, 18*, 103-119. doi: 10.1007/s11065-008-9055-0

Troster, A. I., Fields, J. A., Testa, J. A., Paul, R. H., Blanco, C. R., Hames, K. A.,...Beatty, W. W. (1998). Cortical and subcortical influences on clustering and switching in the performance of verbal fluency tasks. *Neuropsychologia, 36*, 295-304. doi: 10.1016/S0028-3932%2897%2900153-X

Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology, 22*, 370-378.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*, 138-146. doi: 10.1037/0894-4105.11.1.138

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia, 36*, 499-504. doi: 10.1016/S0028-3932%2897%2900152-8

Troyer, A. K., Moscovitch, M, Winocur, G., Leach, L., & Freedman, M. (1998). Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *Journal of the International Neuropsychological Society, 4*, 137-143.

Van Hooren, S. A. H., Valentijn, A. M., Bosma, H., Ponds, R. W. H. M., van Boxtel, M. P. J., & Jollies, J. (2007). Cognitive functioning in healthy older adults aged 64-81: A cohort study into the effects of age, sex, and education. *Aging, Neuropsychology, and Cognition, 14*, 40-54. doi: 10.1080/138255890969483

Wittenberg, D., Possin, K. L., Rascovsky, K., Rankin, K. P., Miller, B. L., & Kramer, J. H. (2008). The early neuropsychological and behavioral characteristics of frontotemporal dementia. *Neuropsychology Review, 18*, 91-102. doi: 10.1007/s11065-008-9056-z

Young, C. J. (2004). Contributions of metaknowledge to retrieval of natural categories in semantic memory. *Journal of Experimental Psychology, 30*, 909-916. doi: 10.1037/0278-7393.30.4.909

Running head: COMPUTER SCORING FOR VERBAL FLUENCY

Computer Scoring Improves Reliability of Calculating Clustering and Switching Rates for Semantic and Phonemic Verbal Fluency Tasks

Nicole Haugrud, Laura Sayer, Craig Thompson, Margaret Crossley, & Mirna Vrbancic
University of Saskatchewan

Author Note

Nicole Haugrud, Department of Psychology, University of Saskatchewan; Laura Sayer, Department of Psychology, University of Saskatchewan; Craig Thompson, Department of Computer Science, University of Saskatchewan; Margaret Crossley, Department of Psychology, University of Saskatchewan; Mirna Vrbancic, Department of Clinical Health Psychology, Royal University Hospital/University of Saskatchewan.

This research was supported in part by a Frederick Banting and Charles Best Canada Graduate Scholarship from the Canadian Institutes of Health Research awarded to Nicole Haugrud.

Correspondence concerning this article should be addressed to Nicole Haugrud, Department of Psychology, Arts 154, University of Saskatchewan, 9 Campus Drive, Saskatoon, Saskatchewan, Canada, S7N 5A5. Email: nicole.haugrud@usask.ca

Abstract

Computerized scoring on neuropsychological tests can improve scoring accuracy and reduce the amount of time required for test interpretation (Woo, 2008). The aim of the current study was to develop a computer scoring program for verbal fluency tasks. Verbal fluency tasks can be scored for total word production and for subcomponents of production including clustering and switching rates (Troyer, Moscovitch, & Winocur, 1997). Computer and hand scoring of semantic and phonemic verbal fluency tasks were compared for 132 healthy individuals. Results showed high consistency between computer and hand scoring for phonemic fluency variables. In contrast to computer scoring, hand scoring for semantic fluency produced a significant number of errors and inconsistencies by well-trained scorers. Additionally, for both semantic and phonemic fluency, computer scoring reduced the time required to calculate dependent measures when compared to hand scoring. Accurate and time-efficient scoring of verbal fluency tasks can contribute to accurate diagnosis in clinical settings. These results indicate that measures of clustering and switching rates for verbal fluency tasks should be calculated using a computer scoring program.

Keywords: computerized assessment, verbal fluency, clustering, switching, neuropsychology

Computer Scoring Improves Reliability of Calculating Clustering and Switching Rates for Semantic and Phonemic Verbal Fluency Tasks

As computer technology continues to advance and become more accessible there is increased potential to integrate computerized assessment and scoring measures into neuropsychology. Computerized neuropsychological test scoring presents a number of potential benefits. Computerization of scoring reduces human error in scoring and increases standardization of test administration (Wild, Howieson, Webbe, Seelye, & Kaye, 2008; Woo, 2008). As well, computerization results in quicker test administration and scoring (Leposvic, Leposavic, & Saula-Marojevic, 2010). Results are available immediately and therefore the time required for scoring is reduced (Woo, 2008). This can result in reduced cost of neuropsychological assessment as fewer hours are required by a trained neuropsychologist and can result in reduced materials costs (Wild et al., 2008). As well, computerized administration simplifies data storage, reduces data entry errors, and makes participant data more easily de-identified (Cernich, Brennana, Barker, & Bleiberg, 2007; Schatz & Zillmer, 2003).

Previous research clearly supports that computerization results in more consistent and reliable scoring of neuropsychological assessment data (Butcher, Perry, & Atlis 2000; Cernich et al., 2007; Wild et al., 2008; Woo, 2008). The aim of the current study was to develop a computerized scoring system for analysis of performance on verbal fluency tasks. Verbal fluency tasks are speeded word generation tasks that require individuals to rapidly produce as many words as possible on a sixty second trial either beginning with a specific letter (e.g. “C”, “F”, “L” on phonemic fluency tasks) or belonging to a specific semantic category (e.g. “animals” on semantic fluency tasks; Strauss, Sherman, & Spreen, 2006). Typically, a score of total words produced is used as a measure of performance on these tasks; however, some researchers have proposed a process approach to interpreting performance on these measures. For example, Troyer and colleagues (1997) proposed calculating measures of clustering (i.e., grouping semantically or phonemically related words) and switching (shifting between clusters of words). On the Animal Naming task (Strauss et al., 2006), individuals might produce a cluster of farm animals (e.g., pig, horse, cow), and then switch to a cluster of pets (e.g., dog, cat, budgie). Other researchers have proposed subdividing switching into hard and cluster switching (Abwender, Swan, Bowerman, & Connolly, 2001). Hard switching is shifting between single word clusters or between a single word and a clustered word, while cluster switching is shifting between two

multiple word clusters (Abwender et al., 2001). Recently, Lanting, Haugrud, and Crossley (2009) proposed examining the number of novel and repeated subcategories accessed, and the percentage of clustered words. Clustering and switching variables have been compared in healthy aging (Haugrud, Lanting, & Crossley, 2010; Lanting et al., 2009) and dementia (Beatty, Testa, English, & Winn, 1997; Epker, Lacritz, & Munro Cullum, 1999; Gomez & White, 2006; Haugrud, Crossley, & Vrbancic, 2011; Troster et al., 1998) research. However, calculation of these variables is time consuming for the researcher and impractical for the clinician, and the calculation of multiple variables can potentially reduce scoring accuracy. The current study describes the development of a computerized scoring program to calculate verbal fluency variables, including measures of clustering and switching. Computer generated scores were compared to scores previously published using hand scoring methods that were checked and re-checked for accuracy (Lanting et al., 2009) to demonstrate the reliability of calculated scores. It was hypothesized that the computer scoring program results would be highly consistent with the hand scoring results reported in previous research (Lanting et al., 2009) and that computer scoring would reduce time required to calculate scores.

Methods

Participants

Participant verbal fluency data were taken from an archival study of healthy aging collected through the Aging, Research, and Memory Clinic in Saskatoon, Saskatchewan. All data were originally collected in compliance with the ethical guidelines of the University of Saskatchewan. Verbal fluency data from 132 participants were used in the current study. Hand scoring results from these participants have previously been reported by Lanting and colleagues (2009). The study sample was comprised of 60 participants in a young adult group ranging in age from 20 to 40 years (M = 28.8, SD = 6.2), and 72 participants in an older adult group ranging in age from 65 to 90 years (M = 74.7, SD = 5.8).

Materials

At the time of initial participation as part of a comprehensive neuropsychological research battery, participants completed the Controlled Oral Word Association Test (COWAT; Benton & Hamsher, 1976) as a measure of phonemic fluency and the Animal Naming test (AN; Strauss et. al, 2006) as a measure of semantic fluency.

Procedures

Hand scoring was completed by Lanting and colleagues (2009) for verbal fluency variables. All fluency variables are calculated for each 60s trial. On the phonemic task, the three trials were added together to produce a phonemic total score for each variable, with the exception of average cluster size where the three trials were averaged. Total word production was calculated for each trial as the number of words generated minus words that are errors and repetitions (Strauss et. al, 2004).

Detailed scoring procedures for the calculation of average cluster size and number of switches have been reported previously (Troyer et al., 1997). Briefly, on the phonemic fluency task, a cluster is a set of sequential words that start with the same two letters, rhyme, differ by only a vowel sound, or are homonyms (Troyer et al., 1997). For example, the words "farm" and "face" would be a cluster of two words on the phonemic task because they start with the same two letters. On the semantic fluency task, a cluster is a group of words that belong to the same semantic subcategory (Troyer et al., 1997). For example, on the Animal Naming task the words "cow" and "horse" would be a cluster of two farm animals. A score of average cluster size is calculated for each trial. In this calculation, a single word is given a score of 0, two clustered words are given a score of 1, and so on. In other words, the size of a cluster equals the number of words in the cluster minus 1. These cluster scores are summed and then divided by the number of clusters in a trial to produce an average cluster size score (Troyer et al., 1997). According to Troyer and colleagues (1997) a switch is a shift between two clusters. The total number of switches on a trial is equal to the number of clusters in a trial minus 1.

Hard and cluster switches have been described in detail by Abwender and colleagues (2001). Briefly, a hard switch is a shift between two single words or between a single word and a clustered word. Each hard switch is given a score of 1 and these hard switches are summed across each trial. A cluster switch is a shift between two multiple word clusters. Each cluster switch is given a score of 1 and these cluster switches are summed across each trial.

Novel and repeated clusters and percentage of clustered words have been described by Lanting and colleagues (2009). A novel cluster on phonemic fluency is a cluster of words that start with the same first two letters that has not been previously used by the participant on that trial. A repeated cluster on phonemic fluency occurs when a participant returns to a previously used cluster to generate new exemplars from that phonemic category. For example, if an

individual produced the words “fast, farm, flip, fly, fake” the words “fast” and “farm” would be a novel cluster starting with “fa”, the words “flip” and “fly” would be a novel cluster starting with “fl” and the word “fake” would be a repeated cluster starting with “fa”. On semantic fluency a novel cluster occurs when clusters of words belong to different semantic subcategory and a repeated cluster occurs when a previously used cluster is returned to later in the same trial (Lanting et al., 2009). For example, on the Animal Naming task if an individual produced the words “cow, horse, lion, monkey, pig” the words “cow” and “horse” would be a novel cluster of Farm Animals, the words “lion” and “monkey” would form a novel cluster of African Animals, and the word “pig” would be a repeated cluster of Farm Animals. The number of multiple word novel clusters and multiple word repeated clusters were also calculated. The same procedure was used as for calculating novel and repeated clusters, however single words were excluded from the analysis and not counted as clusters (Lanting et al., 2009). Finally the percentage of clustered words was calculated by dividing the number of words that belong to a multiple word cluster by the total words produced on each trial (Lanting et al., 2009).

In total 11 variables were calculated for both phonemic and semantic fluency. The average time taken to score each participant’s production on these tasks using had scoring was 5-10 minutes. For participants who produced more atypical words this time requirement for hand scoring increased. In addition, the time required to train individuals in hand scoring procedures was extensive. This training involved reading the scoring procedures, practice scoring for each rater, and comparison of scoring across raters for accuracy. When there were inconsistencies between raters this training was repeated. For the Lanting and colleagues (2009) study this training was conducted over approximately two weeks of daily sessions. Hand scoring in the Lanting and colleagues (2009) study therefore required significant time and resources to complete.

A computer program was developed in collaboration with a graduate student in computer sciences to calculate scores for the verbal fluency variables. The computer program is written in the Python programming language and relies on word lists to group output according to scoring procedures. Consistent with the original scoring procedures of Troyer and colleagues (1997), for both the semantic and phonemic tasks, clusters were generated to maximize cluster size. The computer program starts with the first word in the output and forms the largest possible cluster for that word, then moves to the next word and so on. Consider the following example on

Animal Naming where an individual produces the output "elephant, giraffe, leopard, panther". Starting with the first word, "elephant" belongs to the subcategory "African Animals" (group A). The next word "giraffe" also belongs to the subcategory "African Animals" (group A). "Leopard" could belong to the subcategory "African Animals" or "Feline" (group A or B), as could "panther." Because the computer program categorizes each exemplar sequentially to maximize potential cluster size these four words would be a cluster of 4 African Animals, rather than a cluster of 2 African Animals and 2 Felines. If words generated could be grouped into two categories on the Animal Naming task, and inclusion did not impact cluster size, the superordinate category of living environment would be used. For example if on the Animal Naming task an individual only generated "leopard, panther" this would be recorded as a cluster of two African Animals rather than two Felines because African Animals is the superordinate living environment category.

For phonemic fluency, the computer scoring program was created with a slight modification to the original scoring measures of Troyer and colleagues (1997) who counted clusters as words that shared the same first two letters, rhymed, differed by a vowel sound, or were homonyms. The computer program was able to count as a cluster only words that began with the same two letters. Words that were homonyms, rhymed, or differed by a vowel sound were not counted as clusters. The modification was made because the computer program required to analyze these potential clusters would need to be significantly more complex in order to encode all the possible homonyms, rhyming words, or words that differ by a vowel sound. As a result of this scoring modification, words could not overlap more than one cluster on the phonemic task. However, these phonemic clusters occurred infrequently in the data and therefore it was determined exclusion of these potential groupings was not likely to significantly reduce the reliability of computerized scoring.

For the current study, each participant's raw data were entered into a separate plain text file for each 60 second trial. One participant therefore would have four plain text files (i.e., one for animal names, one for the letter "C", one for the letter "F", and one for the letter "L"). Files were saved according to participant number and task (e.g., for participant number 100 they would have files 100.C, 100.F, 100.L, and 100.animals) which the computer program recognized and analyzed the data according to the specified verbal fluency task. Output from the computer

program is provided in comma separated values format with column headings as variables and rows as participant numbers. Variables can be analyzed with intrusions included or excluded.

For the current study each participant's raw fluency data were entered into plain text format files and then analyzed using the computer program. These data were transferred to SPSS along with the hand scoring data. To assess the reliability of the computer scoring program, intraclass correlation coefficients (ICC) were used to compare hand scoring to computer generated scores for each participant. ICC assesses the agreement between two raters. A value of 1.0 indicates perfect agreement and values from 0.7 to 1.0 indicate high interrater reliability. In contrast to the hand scoring procedure, computer scoring required approximately one minute per participant. This is a significant decline from the average time of 5-10 minutes per participant required to hand score clustering and switching variables. For the current data, hand scoring would have required between 11 hours and 22 hours compared to 2 hours and 12 minutes required for computer scoring and checking.

Results

Phonemic Verbal Fluency

Table 1 gives the means and standard deviations for the phonemic fluency variables scored by hand and by the computer program as well as the intraclass correlation coefficients for computerized versus hand scoring of phonemic fluency variables. On phonemic fluency all variables showed highly significant correlations between computerized and hand scoring.

Table 1

Phonemic Verbal Fluency Variable Means (Standard Deviations) Calculated by Hand and Computer Scoring, and the I*ntraclass Correlations (ICC) Between Hand and Computer Scoring*

Variable	Hand Scoring *M(SD)*	Computer Scoring *M(SD)*	Intraclass Correlation Coefficient *(ICC)*
Total Words Produced[a]	42.4(11.2)	42.3(11.2)	0.992*
Total Switches[b]	28.9(9.1)	28.8(9.0)	0.948*
Average Cluster Size[c]	0.52(0.37)	0.48(0.27)	0.844*
Hard Switches[d]	25.7(8.8)	26.7(9.0)	0.966*
Cluster Switches[e]	2.7(2.6)	2.1(1.9)	0.756*
Novel Clusters[f]	14.2(2.3)	14.1(2.2)	0.929*
Repeated Clusters[g]	17.6(7.8)	17.7(7.7)	0.987*
Multiple Word Novel Clusters[h]	6.5(2.3)	6.4(2.2)	0.962*
Multiple Word Repeated Clusters[i]	2.3(1.8)	2.2(1.8)	0.929*
Percentage of Clustered Words[j]	53.7(14.8)	47.7(13.7)	0.798*

[a]Total words produced is scored as the sum of all words on a 60 second trial minus errors and perseverations; the three phonemic trials are summed to produce a total phonemic score. [b]Total switches is scored as the number of clusters of words on a 60 second trial minus one; the three phonemic trials are summed to produce a total phonemic score. [c]Average cluster size is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one; the three phonemic trials are averaged to produce a total average phonemic score. [d]Hard switches is the sum of switches between two single words or between a single word and a clustered word on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [e]Cluster switches is scored as the sum of switches between two clusters of more than one word on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [f]Novel clusters is scored as the sum of new phonemic subcategories accessed during a 60 second trial; the three phonemic trials are summed for a total phonemic score. [g]Repeated clusters is scored as the sum of all phonemic subcategories an individual returns to on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [h]Multiple word novel clusters is scored as the sum of new phonemic subcategories accessed during a 60 second trial where clusters of single words are excluded from the analysis; the three

phonemic trials are summed for a total phonemic score. [i]Multiple word repeated clusters is scored as the sum of all phonemic subcategories an individual returns to on a 60 second trial where clusters of single words are excluded from the analysis; the three phonemic trials are summed for a total phonemic score. [j]Percentage of clustered words is the percent of the total words produced on a 60 second trial that are grouped in multiple word clusters; the three phonemic trials are averaged for a total phonemic score.

* indicates high level of correlation for ICC

Semantic Verbal Fluency

Table 2 gives the means and standard deviations for hand and computer scoring on semantic verbal fluency as well as the intraclass correlation coefficients for computerized versus hand scoring of semantic fluency variables. The correlations for semantic fluency were smaller in size but scoring was significantly correlated for all semantic fluency variables except average cluster size, multiple word repeated clusters, and percentage of clustered words.

Table 2

Semantic Verbal Fluency Variable Means (Standard Deviations) Calculated by Hand and Computer Scoring, and the I*ntraclass Correlations (ICC) Between Hand and Computer Scoring*

Variable	Hand Scoring *M(SD)*	Computer Scoring *M(SD)*	Intraclass Correlation Coefficient *(ICC)*
Total Words Produced[a]	22.6(6.2)	22.3(6.2)	0.989*
Total Switches[b]	9.6(3.7)	11.0(4.0)	0.838*
Average Cluster Size[b]	1.44(0.88)	1.13(0.58)	0.690
Hard Switches[d]	6.2(4.0)	8.0(4.2)	0.734*
Cluster Switches[e]	3.4(2.0)	2.9(1.8)	0.763*
Novel Clusters[f]	7.1(2.1)	7.9(2.2)	0.702*
Repeated Clusters[g]	3.4(2.2)	4.1(2.5)	0.792*
Multiple Word Novel Clusters[h]	5.1(1.6)	5.1(1.6)	0.868*
Multiple Word Repeated Clusters[i]	1.0(1.0)	0.9(0.9)	0.649
Percentage of Clustered Words[j]	83.4(19.2)	74.2(15.0)	0.540

[a]Total words produced is scored as the sum of all words on a 60 second trial minus errors and perseverations. [b]Total switches is scored as the number of clusters of words on a 60 second trial minus one. [c]Average cluster size is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one. [d]Hard switches is the sum of switches between two single words or between a single word and a clustered word on a 60 second trial. [e]Cluster switches is scored as the sum of switches between two clusters of more than one word on a 60 second trial. [f]Novel clusters is scored as the sum of new animal subcategories accessed during a 60 second trial. [g]Repeated clusters is scored as the sum of all animal subcategories an individual returns to on a 60 second trial. [h]Multiple word novel clusters is scored as the sum of new animal subcategories accessed during a 60 second trial where single word clusters are excluded from the analysis. [i]Multiple word repeated clusters is scored as the sum of all animal subcategories an individual returns to on a 60 second trial where single word clusters are excluded from the analysis. [j]Percentage of clustered words is scored as the percentage of the total words produced on a 60 second trial that are grouped into multiple word clusters.

* indicates high level of correlation for ICC

Lanting and Colleagues (2009) Reanalysis

Because the correlations between hand and computer scoring were not perfect, and were low for some semantic fluency variables, the comparisons between the young and older age group conducted in the study by Lanting and colleagues (2009) were re-analyzed using the computerized scoring results. Age group (2) by sex (2) analyses of variance were run for each phonemic and semantic verbal fluency variable. Partial eta squared was used as a measure of effect size.

Phonemic fluency age group differences.

When the analysis was run for the phonemic fluency variables using computer generated scores, age effects were consistent with those reported by Lanting and colleagues (2009) for all variables except novel clusters. Using computer generated results, the main effect of age was significant for phonemic novel clusters, $F(1,128) = 5.944$, $p = .016$, $\eta_p^2 = .044$, with the young age group producing more novel clusters than the older age group. In the Lanting and colleagues (2009) results this effect approached significance ($p = .083$) and in the current analysis the effect is small ($\eta_p^2 = .044$) indicating the computer scoring result is consistent with hand scoring.

Phonemic fluency sex effects.

Phonemic fluency sex effects were similar to those reported by Lanting and colleagues (2009) for all phonemic fluency variables except percentage of clustered words. Lanting and colleagues (2009) reported a significant sex difference on this variable in favour of men producing more clustered words. With computer scoring, however, this effect only approached significance, $F(1,128) = 3.184$, $p = .077$, $\eta_p^2 = .024$. This effect is also a small effect. This indicates that, even with small differences in significance level observed, the hand and computer scoring results are consistent.

Semantic fluency age group differences.

Semantic fluency age group differences were consistent with those reported by Lanting and colleagues (2009) for all semantic fluency variables except average cluster size and percentage of clustered words. For semantic fluency average cluster size the age group difference was no longer significant when computer scoring was used, $F(1,128) = 1.829$, $p = .179$, $\eta_p^2 = .014$. Intraclass correlations showed a large degree of discrepancy between hand and computer scoring for semantic fluency average cluster size due to the difficulty in hand scoring this variable. As well, the age effect reported by Lanting and colleagues (2009) showed a small effect

size (η_p^2 = .064). The difficulty hand scoring semantic fluency average cluster size and the small observed age effect likely resulted in a difference in observed significance between computer and hand scoring. Computer scoring also eliminated the age group difference for percentage of clustered words, $F(1,128) = 2.356$, $p = .127$, $\eta_p^2 = .018$, reported by Lanting and colleagues (2009). Similarly to average cluster size, semantic fluency percentage of clustered words showed poor interrater reliability between hand and computer scoring and the age effect reported on this variable by Lanting and colleagues (2009) was small (η_p^2 = .034), which likely accounts for this discrepancy.

Semantic fluency sex effects.

Using computerized scoring, observed sex effects on semantic fluency reported by Lanting and colleagues (2009) were no longer significant for average cluster size, $F(1,128) = 2.434$, $p = .121$, $\eta_p^2 = .019$, novel clusters, $F(1,128) = 1.472$, $p = .227$, $\eta_p^2 = .011$, or percentage of clustered words, $F(1,128) = 1.216$, $p = .272$, $\eta_p^2 = .009$. These variables showed poor consistency between hand and computer scoring. In addition, the sex effects reported by Lanting and colleagues (2009) on these variables were small. Inconsistencies due to difficulties in hand scoring and small effect sizes likely produced differences in observed sex effects between hand and computer scoring. As well, observed interaction effects reported by Lanting and colleagues (2009) were no longer significant for semantic fluency hard switches, $F(1,128) = 1.278$, $p = .260$, $\eta_p^2 = 010$, or cluster switches, $F(1,128) = 2.206$, $p = .140$, $\eta_p^2 = .017$. The interaction effects previously reported using hand scoring were small on these variables, indicating slight differences between hand and computer scoring might have eliminated observed effects.

Discussion

Consistent with the study hypothesis, the computer scoring of the phonemic fluency variables was highly correlated with the hand scoring results published by Lanting and colleagues (2009). The only phonemic fluency variables where correlations between hand and computer scoring were less than $r = 0.900$ were average cluster size, cluster switches, and percentage of clustered words. The computer program used a minor alteration to the scoring of phonemic clustering compared to the original hand scoring measures described by Troyer and colleagues (1997). For the computer scoring, phonemic clusters were only possible if two consecutive words started with the same first two letters. Due to the complexity required to produce a computerized scoring program that could cluster words that were homonyms, words

that rhyme, or words that differ only by a vowel sound these potential clusters were excluded from the computer program. The results of the current study showed a high level of consistency between phonemic hand and computer scoring, indicating this small change did not significantly impact results. Another potential reason for non-perfect correlations between hand and computer scoring on phonemic fluency is potential human error in hand scoring. All word entry on the computer program was double checked by an independent data entry person to reduce human error at data entry. However small errors in scoring were observed in the hand scored data on phonemic fluency, producing some variability between the hand and computer scoring. In addition, when the age group by sex analyses from Lanting and colleagues (2009) were re-run using computerized scoring, results were highly consistent with the previously published results. Only two variables, the age group effect on novel clusters and the sex effect on percentage of clustered words showed differences between hand and computer scoring. However, these differences were small and likely related to small effect sizes in the original analysis. Taken together the phonemic fluency results indicate high reliability between hand and computer scoring with small differences due to slight modifications to scoring procedures for the computer program and very occasional and minor human error in hand scoring.

The semantic fluency analysis provided partial support for the study hypothesis. Hand and computer scoring were highly correlated for semantic total words produced. Computerized and hand scoring were also highly correlated for total switches, hard switches, cluster switches, novel clusters, repeated clusters, and multiple word novel clusters. However, these correlations were not as strong as for phonemic fluency. As well, the semantic fluency correlations did not indicate high reliability between hand and computer scoring for average cluster size, multiple word repeated clusters, and percentage of clustered words. Although the computer program used the same scoring procedures as defined by Troyer and colleagues (1997) for average cluster size on semantic fluency, there were observed differences between the hand and computer scoring. For hand scoring of verbal fluency variables, scoring was practiced for consistency between raters before commencing scoring of the study data. As well, a significant portion of the hand scoring was rechecked for accuracy by a second rater. Even with checking of hand scoring on semantic and phonemic fluency, errors were observed in the hand scoring. For example, rules for semantic cluster size calculation were not always consistently applied by hand scorers. Specifically, variability increased when participants generated animals that were less common

exemplars and exemplars that could be used in multiple categories (e.g. “leopard” could be either “African animals” or “feline”). Although the scoring guidelines provided by Troyer and colleagues (1997) stated specifically that the superordinate category should be used first to group semantic clusters, this scoring criteria was not always consistently applied during hand scoring. Computerized scoring eliminated this inconsistency on semantic fluency and provided more accurate semantic fluency scoring. Even with observed inconsistencies between hand and computer scoring on semantic fluency, when the age group by sex analysis reported by Lanting and colleagues (2009) was re-run using computerized scoring the results were largely consistent with hand scoring results. Discrepancies between the age and sex effects observed using the hand scoring of Lanting and colleagues (2009) and the computer scoring in the current study resulted mainly from small effect sizes reported by Lanting and colleagues for some variables. Because these effect sizes are small, minor changes to scoring consistency can produce changes in significance values. As well, the variables where hand and computer scoring differences were larger (i.e. average cluster size and percentage of clustered words) showed the largest effect differences when the age by sex comparisons were re-analyzed. This further supports the use of the more accurate and consistent computer scoring program.

A second hypothesis of the current study was that computerized scoring would be more time efficient than hand scoring. Hand scoring the verbal fluency measures assessed in this study required approximately 5-10 minutes per participant to score all four subtests (animals, C, F, and L). Computerized scoring reduced that time to approximately 1 minute per participant to enter the words into the computer. Running the computer program required only a few seconds. This is a 4-9 minute per participant reduction in the scoring time required using the computer program for a total time saved of 8.8-19.8 hours through the use of the computer program in this study. This time savings, combined with more accurate and consistent scoring, indicates the use of a computer scoring program for verbal fluency tasks could be a valuable contribution to both clinical work and research in this area.

Taken together the results of the current study indicate the computer program developed to score clustering and switching variables on verbal fluency tasks provides more efficient and more accurate scoring of these variables than hand scoring procedures. This is consistent with previous research which has shown computerized neuropsychological assessment improves scoring accuracy and speed of assessment (Butcher et al., 2000; Cernich et al., 2007; Wild et al.,

2008; Woo, 2008). Verbal fluency tasks are frequently used in clinical settings to detect cognitive impairment. Quicker and more accurate scoring of assessment results is likely to improve diagnostic decision making and the generation of treatment recommendations. A limitation of the current study is that verbal fluency was only assessed in a healthy adult group. Individuals with cognitive impairment such as dementia tend to produce more errors (Azuma, 2004) which could make it more difficult for the computer program to accurately categorize responses. Each error term would have to be added to the computer program scoring template which might potentially increase the time required to use this program with a clinical sample. Future research should validate the use of this computer program with a clinical sample.

References

Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment, 8*, 323-336.

Azuma, T. (2004). Working memory and perseveration in verbal fluency. *Neuropsychology, 18*, 69-77.

Beatty, W. W., Testa, J. A., English, S., & Winn, P. (1997). Influences of clustering and switching on the verbal fluency performance of patients with Alzheimer's disease. *Aging, Neuropsychology, and Cognition, 4*, 273-279.

Butcher, J. N., Perry, J. N., & Atlis, M. M. (2000). Validity and utility of computer-based test interpretation. *Psychological Assessment, 12*, 6-18. doi: 10.1037/1040-3590.12.16.

Cernich, A. N., Brennana, D. M., Barker, L. M., & Bleiberg, J. (2007). Sources of error in computerized neuropsychological assessment. *Archives of Clinical Neuropsychology, 22S*, S39-S48. doi: 10.1016/j.acn.2006.10.004.

Epker, M. O., Lacritz, L. H., & Munro Cullum, C. (1999). Comparative analysis of qualitative verbal fluency performance in normal elderly and demented populations. *Journal of Clinical and Experimental Neuropsychology, 21*, 425-434.

Gomez, R. G., & White, D. A. (2006). Using verbal fluency to detect very mild dementia of the Alzheimer type. *Archives of Clinical Neuropsychology, 21*, 771-775. doi: 10.1016/j.acn.2006.06.012

Haugrud, N., Lanting, S., & Crossley, M. (2010). The effects of age, sex and Alzheimer's disease on strategy use during verbal fluency tasks. *Aging, Neuropsychology, & Cognition, 17*. 220-239. doi: 10.1080/13825580903042700.

Haugrud, N., Crossley, M., & Vrbacic, M. (2011). Clustering and switching strategies during verbal fluency performance differentiate Alzheimer's disease and healthy aging. *Journal of the International Neuropsychological Society, 17*, 1153-1157. doi: 10.1017/S1355617711001196.

Lanting, S., Haugrud, N., & Crossley, M. (2009). The effects of age and sex on clustering and switching during speeded verbal fluency tasks. *Journal of the International Neuropsychological Society, 15*, 196-204. doi: 10.1017/S1355617709090237

Schatz, P., Zillmer, E. A. (2003). Computer-based assessment of sports-related concussion. *Applied Neuropsychology, 10*, 42-47.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary, 3rd Ed.*. New York: Oxford University Press.

Troster, A. I., Fields, J. A., Testa, J. A., Paul, R. H., Blanco, C. R., Hames, K. A.,...Beatty, W. W. (1998). Cortical and subcortical influences on clustering and switching in the performance of verbal fluency tasks. *Neuropsychologia, 36*, 295-304. doi: 10.1016/S0028-3932%2897%2900153-X

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*, 138-146. doi: 10.1037/0894-4105.11.1.138

Wild, K., Howieson, D, Webbe, F., Seelye, A., Kaye, J. (2008). Status of computerized cognitive testing in aging: A systematic review. *Alzheimer's & Dementia, 4*, 428-437. doi: 10.1016/j.jalz.2008.07.003.

Woo, E. (2008). Computerized neuropsychological assessments. *CNS Spectrums, 13* (Suppl 16), 14-17.

Running head: FLUENCY SUBCOMPONENTS AND HEALTHY AGING

Analysis of the Subcomponents of Verbal Fluency Production Supports the Frontal Executive Hypothesis of Healthy Aging

Nicole Haugrud, Margaret Crossley, & Mirna Vrbancic
University of Saskatchewan

Author Note

Nicole Haugrud, Department of Psychology, University of Saskatchewan; Margaret Crossley, Department of Psychology, University of Saskatchewan; Mirna Vrbancic, Department of Clinical Health Psychology, Royal University Hospital.

This research was supported in part by a Frederick Banting and Charles Best Canada Graduate Scholarship from the Canadian Institutes of Health Research awarded to Nicole Haugrud. The authors wish to acknowledge Michelle Shaw for her contribution to data collection.

Correspondence concerning this article should be addressed to Nicole Haugrud, Department of Psychology, Arts 154, University of Saskatchewan, 9 Campus Drive, Saskatoon, Saskatchewan, Canada, S7N 5A5. Email: nicole.haugrud@usask.ca

Abstract

Verbal fluency tasks require individuals to rapidly produce exemplars of a given semantic or phonemic category. Intact processing speed, mental set shifting, search and retrieval abilities, and memory storage are required to perform well on these tasks (Henry & Crawford, 2004). Given the number of cognitive components required to complete these tasks, examining subcomponents of verbal fluency performance beyond total word production is informative for both clinical and healthy aging research. Two subcomponents proposed by Troyer et al. (1997) are clustering (i.e., generating groups of semantically or phonemically related words) and switching (i.e., shifting between clusters). The current study examined measures of clustering and switching in 90 healthy adults divided into young (20-38 yrs), middle-aged (40-63 yrs), and older (65-82 yrs) groups. The older age group produced fewer semantic but equivalent phonemic total words when compared to middle-aged and younger groups, and fewer hard switches (i.e., shifts between single-word clusters; Abwender et al., 2001) on both fluency tasks. There were no age group differences for average cluster size. These results are consistent with age related declines in processing speed and mental set shifting, and age-related stability for memory storage (Bryan & Luszcz, 2000). This study supports the frontal executive hypothesis of healthy aging (MacPherson et al., 2002) and demonstrates the value of examining specific components of verbal fluency performance such as clustering and switching strategies.

Keywords: verbal fluency, clustering, switching, aging, executive functions

Analysis of the Subcomponents of Verbal Fluency Production Supports the Frontal Executive Hypothesis of Healthy Aging

Verbal fluency tasks are speeded word generation tasks that require participants to either list words beginning with a specific letter (phonemic fluency) or belonging to a specific category (semantic fluency). Optimal semantic and phonemic verbal fluency performance is presumed to be reliant on healthy temporal and prefrontal lobe functioning, respectively (Mummery, Patterson, Hodges, & Wise, 1996). A score of total words produced in a 60-second trial is the most commonly used performance measure.

Age related differences on verbal fluency tasks have been previously examined, with some inconsistencies in results. Older age groups typically produce fewer words on semantic fluency tests compared to younger groups (Bryan & Luszcz, 2000; Clark et al., 2009; Crossley, D'Arcy, & Rawson, 1997; Kavé, 2005; Lanting, Haugrud, & Crossley, 2009). On phonemic fluency tasks, some studies report no differences in performance across age (Bryan & Luszcz, 2000; Crossley et al., 1997), whereas other studies have reported significant age-related differences, although the age effects typically are smaller than for semantic fluency (Brickman et al., 2005; Lanting et al., 2009; Rodriguez-Aranda & Martinussen, 2006).

An alternative approach is to compare the specific task components required for normal performance on verbal fluency tasks. For example, although phonemic fluency is generally thought to be relatively more dependent on intact prefrontal lobe functioning and semantic fluency is presumed to be relatively more dependent on intact temporal lobe functioning (Mummery et al., 1996), both tasks require verbal abilities, search and retrieval skills, adequate speed of processing, and an ability to inhibit inappropriate responses (Abwender, Swan, Bowerman, & Connolly, 2001; Henry & Crawford, 2004). The two component model of verbal fluency described by Troyer, Moscovitch, and Winocur (1997) is an example of a process approach that goes beyond total word production. Troyer and her colleagues proposed that verbal fluency performance requires both the production of words within either a semantic or phonemic subcategory (i.e., clustering) and the ability to shift between clusters (i.e., switching). Clustering is presumed to rely on temporal lobe functions to produce exemplars of either phonemic or semantic categories, and switching is presumed to rely on prefrontal lobe functions for strategic search processes, and these distinctions are supported by lesion and imaging research (Hirshorn & Thompson-Schill, 2006; Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998).

Age group differences during verbal fluency tasks have been examined using measures of clustering and switching. A number of researchers have found that older adults switch less frequently on verbal fluency tasks when compared to younger groups (Bruicki & Rocka, 2004; Haugrud, Lanting, & Crossley, 2010; Lanting et al., 2009; Troyer et al., 1997; 2000). Since aging effects have long been associated with decreases in executive functioning, this finding is consistent with the frontal executive hypothesis of healthy aging (Henry & Phillips, 2006; MacPherson, Phillips, & Sala, 2002) and, additionally, provides support for a two component model of verbal fluency. In contrast to the switching data, some have reported that older adults produce larger phonemic clusters (Troyer et al., 1997; 2000) or larger semantic clusters (Lanting et al., 2009) than younger age groups while other studies have reported no age group differences for phonemic or semantic cluster size (Haugrud et al., 2010; Hughes & Bryan, 2002). In studies that have found a clustering advantage for older adults these effects have tended to be small (Lanting et al., 2009; Troyer et al., 1997; 2000). As well some of these previous studies have described older adults with atypically high education levels, which could artificially advantage older adults on verbal fluency by creating cohort differences in general verbal ability (Crossley et al., 1997). Perhaps as a result of these sampling differences, reports of the effect of age on clustering data have been inconsistent.

Although Troyer and colleagues (1997) proposed measuring clustering and switching as a method to specify the cognitive abilities required to perform verbal fluency tasks, there are limitations associated with their scoring procedures. For example, Mayer (2002) noted that switching rates in the Troyer and colleagues (1997) model can be impaired either because the individual has difficulties accessing new semantic clusters or because they have difficulty generating words within clusters. To address this concern, Abwender and colleagues (2001) proposed two types of switches (i.e., hard switches and cluster switches). Hard switching occurs between two single words or between a clustered word and a single word, and is believed to reflect the speeded nature of verbal fluency tasks. Cluster switching occurs between two groups of multiple word clusters and is believed to reflect mental flexibility. Lanting and colleagues (2009) examined the number of novel and repeated subcategories accessed and found that older compared to younger adults produced fewer of both. These authors concluded that the generation of new words from repeated clusters is an efficient verbal fluency strategy that is characteristic of younger adults.

Haugrud and colleagues (2010) also proposed modifications to the Troyer and colleagues (1997) method. Troyer and colleagues (1997) included perseverations and errors in the calculation of clustering and switching scores because they presumed that these intrusions might be strategic and prompt individuals to initiate new clusters. If perseverations and errors are not systematic but rather are randomly distributed, then including these intrusions in the measurement of clustering and switching tasks might bias the results. Specifically, this scoring inclusion bias might artificially increase the cluster size and switching scores for older individuals who tend to produce more errors and perseverations than healthy younger individuals. In keeping with these hypotheses, Haugrud and colleagues (2010) found that intrusions were randomly distributed, and that their removal from the calculation of clustering and switching variables lowered the scores on these variables in a study comparing Alzheimer disease participants to a healthy control group, and produced results more consistent with the two component model of verbal fluency.

The current study compared the methods of scoring proposed by Troyer et al. (1997), Abwender et al. (2001), and Lanting et al. (2009) in the investigation of verbal fluency in young, middle aged, and older groups. In addition, this study calculated scores both with errors and perseverations included and excluded as described by Haugrud et al. (2010), and computed average cluster size both with and without single word clusters, to address the concern expressed by previous researchers that a cluster of one does not reflect semantic or phonemic grouping (Lanting et al., 2009).

Based on the two component model of verbal fluency and previous research, it was hypothesized that for both fluency tasks, when compared to the young and middle-age groups, the oldest age group would produce fewer words and switches. In addition, it was hypothesized that the older age group would produce fewer hard switches than the middle and young age groups, due to age related declines in processing speed, but there would be no age effect on cluster switches. No age category effects were hypothesized for average cluster size (regardless of the method used to calculate average cluster size) or for percentage of clustered words. Consistent with a presumed age-related deficit in the search and retrieval process and with Lanting and colleagues' (2009) recent findings, older age groups were hypothesized to produce fewer novel and repeated clusters when compared to younger participants.

Methods

Participants

The current study used archival data from 90 participants recruited for a larger neuropsychological investigation of normal aging. The data included in this manuscript were obtained in compliance with the ethics regulations of the authors' institution. Participants for the current study ranged from 20-82 years of age and were divided into a young age group (n = 30, aged 20-38 years), a middle age group (n = 30, aged 40-63 years), and an older age group (n = 30, aged 65-82). Demographic data for the three age groups is presented in Table 1. Peabody Picture Vocabulary Test-Revised (PPVT-R: Dunn & Dunn, 1981) scores were used as an estimate of verbal intelligence and demonstrated age-equivalence, $F(2,87) = 1.114$, $p = .333$, $\eta_p^2 = .025$.

Table 1

Demographic Data Means (Standard Deviations) for the Young, Middle, and Older Age Group

	Young	Middle	Older
	M(SD)	*M(SD)*	*M(SD)*
Age	27.9(6.0)	51.1(7.6)	71.3(5.9)
Years of Education	15.7(2.6)	14.6(2.9)	13.0(3.5)
PPVT-R Score	161.03(7.72)	164.30(7.97)	162.50(9.64)

Note. PPVT-R is the Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1981).

Materials

As part of a comprehensive neuropsychological research battery, participants completed the Controlled Oral Word Association Test (COWAT; Benton & Hamsher, 1976) as a measure of phonemic fluency and the Animal Naming test (AN; Strauss, Sherman, & Spreen, 2006) as a measure of semantic fluency. As described above, participants completed the Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn & Dunn, 1981) as a measure of verbal intelligence.

Procedures and Scoring

The COWAT consists of three one-minute trials during which participants are required to produce as many words as possible that begin with the letters "C", "F", or "L." On the AN test, participants are given one minute to orally produce as many animal names as possible.

Based on previous research, twelve scores were obtained for each verbal fluency measure: total words produced; average cluster size (calculated both with single words included

and excluded); number of switches; number of hard switches; number of cluster switches; number of novel clusters and number of repeated clusters (calculated both with single words included and excluded); percentage of clustered words; and number of total intrusions. All verbal fluency variables were calculated both with intrusions (i.e., errors and perseverations) included and excluded. On the phonemic task, the three trials were added together to produce a phonemic total score for each variable, with the exception of average cluster size where the three trials were averaged. Detailed scoring procedures for the calculation of average cluster size and number of switches have been reported previously (Troyer et al., 1997; 2000) as have the procedures for calculating hard and cluster switches (Abwender et al., 2001; Lanting et al., 2009). Briefly, a cluster is a set of phonemically or semantically related words (on the phonemic or semantic task, respectively). Hard switches occur between two single words or between a single word and clustered words, and cluster switches occur between two groups of clustered words.

For the current study, a computer program was developed to calculate the verbal fluency scores and to increase the reliability of the scoring procedures. The computer program is written in Python programming language and relies on word lists to group output according to scoring procedures. This program was created with a slight modification to the original scoring method of Troyer and colleagues (1997). According to the original method, a phonemic cluster occurs when successively generated words start with the same first two letters, rhyme, differ only by a vowel sound, or are homonyms. For the current study, on the phonemic task, only the criterion of the same first two letters was used as a cluster. As a result, words could not overlap more than one cluster on the phonemic task. Using this computer scoring method, the verbal fluency scores obtained were consistent with those obtained in previous studies (Haugrud et al., 2010; Lanting et al., 2009; Troyer, 2000), indicating the modification made to the scoring procedures had minimal impact on scores.

The remaining variables have been described in detail by Lanting and colleagues (2009). For the calculation of novel and repeated clusters, clusters were defined by the criteria of Troyer and colleagues (1997) for the semantic tasks, and included words with the same first two letters for the phonemic task. For the semantic task, the superordinate category of living environments was used when a word could be clustered into two different categories, as described by Troyer

and colleagues (1997). Novel and repeated clusters were calculated both including single words as clusters and by excluding single words. Percentage of clustered words was also calculated.

Results

For semantic and phonemic fluency tasks, separate analyses of variance (ANOVAs) were performed for each of the verbal fluency variables. When errors and perseverations were removed from the calculation of verbal fluency variables, effect sizes for the significant findings were larger and consistent with the hypotheses of the current study and past research. Consequently, the following results are presented with intrusions (i.e., errors and perseverations) removed. Partial η_p^2 is reported as a measure of effect size.

Semantic fluency

Refer to Table 2 for the means and standard deviations of the semantic verbal fluency scores according to age group.

Table 2

Semantic Verbal Fluency Means (Standard Deviations) for Young, Middle, and Older Age Groups

	Young	Middle	Older
Total Words Produced[a]	22.7 (4.3)	21.8 (3.5)	17.4 (4.7)***
Number of Switches[b]	10.9 (2.9)	11.0 (2.3)	8.0 (3.0)***
Average Cluster Size (ACS)[c]	1.00 (0.35)	0.91 (0.42)	1.00 (0.39)
ACS (Single Words Removed)[d]	1.95 (0.58)	1.73 (0.49)	1.96 (0.73)
Hard Switches[e]	8.6 (3.8)	8.6 (3.0)	6.2 (3.3)**
Cluster Switches[f]	2.3 (1.6)	2.4 (1.4)	1.8 (1.5)
Novel Clusters[g]	7.8 (1.5)	8.3 (1.3)	6.4 (1.7)***
Repeated Clusters[h]	4.1 (2.5)	3.7 (2.0)	2.6 (2.0)*
Multiple Word Novel Clusters[i]	4.8 (1.3)	5.1 (1.2)	4.0 (1.6)*
Multiple Word Repeated Clusters[j]	0.9 (0.9)	0.56 (0.7)	0.4 (0.6)*
Percentage Clustered Words[k]	72.7(11.0)	70.0(15.7)	72.2(15.5)
Total Intrusions[l]	0.8(1.1)	0.6(1.0)	0.7(0.8)

[a]Total words produced is scored as the sum of all words on a 60 second trial minus errors and perseverations. [b]Total switches is scored as the number of clusters of words on a 60 second trial minus one. [c]Average cluster size is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one. [d]Average cluster size single words removed is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one and single word clusters are excluded from the analysis. [e]Hard switches is the sum of switches between two single words or between a single word and a clustered word on a 60 second trial. [f]Cluster switches is scored as the sum of switches between two clusters of more than one word on a 60 second trial. [g]Novel clusters is scored as the sum of new animal subcategories accessed during a 60 second trial. [h]Repeated clusters is scored as the sum of all animal subcategories an individual returns to on a 60 second trial. [i]Multiple word novel clusters is scored as the sum of new animal subcategories accessed during a 60 second trial where single word clusters are excluded from the analysis. [j]Multiple word repeated clusters is scored as the sum of all animal subcategories an individual returns to on a 60 second trial where single word clusters are excluded from the

analysis. [k]Percentage of clustered words is scored as the percentage of the total words produced on a 60 second trial that are grouped into multiple word clusters. [l]Total intrusions is scored as the sum of all errors and perseverations across a trial

* $p < .05$, ** $p < .01$, *** $p < .001$

Traditional scoring methods.

For semantic fluency total words produced there was a significant main effect of age, $F(2,87) = 13.788$, $p < .001$, $\eta_p^2 = .241$. Analysis of main effects indicates the young and middle age groups produced significantly more words than the older age group. There was no age group difference for number of intrusions produced, $F(2,87) = 0,549$, $p = .579$, $\eta_p^2 = .012$.

Troyer and colleagues (1997) scoring methods.

On semantic fluency number of switches there was a significant main effect of age category, $F(2,87) = 10.913$, $p < .001$, $\eta_p^2 = .201$. The older age group produced fewer switches than both the young and middle age groups. There was no significant main effect of age on semantic fluency average cluster size, $F(2,87) = 0.605$, $p = .548$, $\eta_p^2 = .014$, or on average cluster size with single words excluded, $F(2,87) = 1.404$, $p = .251$, $\eta_p^2 = .031$.

Abwender and colleagues (2001) scoring methods.

There was a significant main effect of age on semantic fluency hard switches, $F(2,87) = 4.974$, $p = .009$, $\eta_p^2 = .103$, with the older age group producing fewer hard switches than the young or middle age groups. The main effect of age on cluster switches was not significant, $F(2,87) = 1.273$, $p < .285$, $\eta_p^2 = .028$.

Lanting and colleagues (2009) scoring methods.

On semantic fluency novel clusters there was a significant main effect of age, $F(2,87) = 12.158$, $p < .001$, $\eta_p^2 = .218$. The older age group produced fewer novel clusters than the young or middle age groups. There was also a significant main effect of age category on semantic fluency repeated clusters, $F(2,87) = 3.642$, $p = .030$, $\eta_p^2 = .077$. The older age group produced fewer repeated clusters than the young age group. The main effect for age on semantic fluency number of multiple word novel clusters also was significant, $F(2,87) = 4.334$, $p = .016$, $\eta_p^2 = .091$, with the older age group producing fewer multiple word novel clusters than the middle age group and the young group. On semantic fluency multiple word repeated clusters there was a significant main effect of age, $F(2,87) = 3.853$, $p = .025$, $\eta_p^2 = .081$, with the older age group producing fewer multiple word repeated clusters than the young age group. There was no significant main effect of age on percentage of clustered words, $F(2,87) = 0.318$, $p = .728$, $\eta_p^2 = .007$.

Phonemic fluency

Refer to Table 3 for the means and standard deviations of the phonemic verbal fluency scores according to age group.

Table 3

Phonemic Verbal Fluency Variable Means (Standard Deviations) for Young, Middle, and Older Age Groups

	Young	Middle	Older
Total Words Produced[a]	40.5 (8.7)	39.8 (11.5)	37.6 (10.4)
Number of Switches[b]	27.2 (7.0)	25.0 (7.5)	22.7 (7.2)*
Average Cluster Size (ACS)[c]	0.38 (0.18)	0.51 (0.32)	0.53 (0.33)
ACS (Single Words Removed)[d]	1.43 (0.53)	1.37 (0.52)	1.64 (0.76)
Hard Switches[e]	25.7 (7.2)	22.9 (7.2)	20.8 (7.4)**
Cluster Switches[f]	1.5 (1.4)	2.2 (1.7)	1.9 (1.7)
Novel Clusters[g]	13.9 (1.9)	13.3 (2.8)	12.5 (2.5)
Repeated Clusters[h]	16.3 (5.8)	14.7 (5.4)	13.2 (5.4)
Multiple Word Novel Clusters[i]	5.3 (2.1)	6.0 (2.0)	5.4 (2.2)
Multiple Word Repeated Clusters[j]	1.9 (1.3)	2.0 (1.9)	1.6 (1.6)
Percentage Clustered Words[k]	43.0(13.0)	47.8(2.5)	48.6(15.3)
Total Intrusions[l]	1.1(1.3)	1.1(1.2)	2.0(2.6)

[a]Total words produced is scored as the sum of all words on a 60 second trial minus errors and perseverations; the three phonemic trials are summed to produce a total phonemic score. [b]Total switches is scored as the number of clusters of words on a 60 second trial minus one; the three phonemic trials are summed to produce a total phonemic score. [c]Average cluster size is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one; the three phonemic trials are averaged to produce a total average phonemic score. [d]Average cluster size single words removed is scored as the averae of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one and single word clusters are removed from the analysis; the three phonemic trials are averaged to produce a total average phonemic score. [e]Hard switches is the sum of switches between two single words or between a single word and a clustered word on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [f]Cluster switches is scored as the sum of switches between two clusters of more than one word on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [g]Novel clusters is scored as the sum of new phonemic subcategories accessed during a 60 second trial; the three phonemic

trials are summed for a total phonemic score. [h]Repeated clusters is scored as the sum of all phonemic subcategories an individual returns to on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [i]Multiple word novel clusters is scored as the sum of new phonemic subcategories accessed during a 60 second trial where clusters of single words are excluded from the analysis; the three phonemic trials are summed for a total phonemic score. [j]Multiple word repeated clusters is scored as the sum of all phonemic subcategories an individual returns to on a 60 second trial where clusters of single words are excluded from the analysis; the three phonemic trials are summed for a total phonemic score. [k]Percentage of clustered words is the percent of the total words produced on a 60 second trial that are grouped in multiple word clusters; the three phonemic trials are averaged for a total phonemic score. [l]Total intrusions is scored as the sum of all errors and repetitions across a 60 second trial; the three phonemic trials are summed for a total phonemic score.

* *p* marginally significant, ** $p < .05$

Traditional scoring methods.

There was no significant effect of age category on phonemic fluency total words produced, $F(2,87) = 0.651$, $p = .524$, $\eta_p^2 = .015$, or total intrusions produced, $F(2,87) = 2.177$, $p = .120$, $\eta_p^2 = .048$.

Troyer and colleagues (1997) scoring methods.

There was no significant main effect of age on phonemic average cluster size, $F(2,87) = 2.453$, $p = .092$, $\eta_p^2 = .053$ or average cluster size excluding single words, $F(2,87) = 1.621$, $p = .204$, $\eta_p^2 = .036$. The main effect of age approached significance on phonemic fluency number of switches, $F(2,87) = 2.811$, $p = .066$, $\eta_p^2 = .061$, with the older age group producing the fewest switches, followed by the middle age group and the young age group.

Abwender and colleagues (2001) scoring methods.

The main effect of age was significant on phonemic fluency number of hard switches, $F(2,87) = 3.360$, $p = .039$, $\eta_p^2 = .072$. The older age group produced significantly fewer hard switches than the young age group. There was no significant main effect of age on phonemic fluency cluster switches, $F(2,87) = 1.310$, $p = .275$, $\eta_p^2 = .029$.

Lanting and colleagues (2009) scoring methods.

The main effect of age was not significant for phonemic fluency novel clusters, $F(2,87) = 2.286$, $p = .108$, $\eta_p^2 = .050$, repeated clusters, $F(2,87) = 2.367$, $p = .100$, $\eta_p^2 = .052$, multiple word novel clusters, $F(2,87) = 1.020$, $p = .365$, $\eta_p^2 = .023$, multiple word repeated clusters, $F(2,87) = 0.557$, $p = .575$, $\eta_p^2 = .013$, or percentage of clustered words, $F(2,87) = 1.435$, $p = .244$, $\eta_p^2 = .032$.

Discussion

The goal of the current study was to examine subcomponents of total word production on verbal fluency tasks in a healthy aging sample in order to investigate which subcomponents show age related change and which demonstrate age stability. Consistent with the hypotheses of the current study, the older age group produced fewer total words on the semantic task. However, we found age-equivalency for phonemic fluency total words. Similarly, previous healthy aging research using phonemic fluency has either found relatively small effect sizes (Brickman et al., 2005; Haugrud et al., 2010; Lanting et al., 2009; Rodriguez-Aranda & Martinussen, 2006) or no age effect for phonemic total word production (Bryan & Luszcz, 2000; Crossley et al., 1997). Because phonemic fluency is believed to be relatively dependent on intact prefrontal lobe

functioning (Mummery et al., 1996) we would expect phonemic fluency to show age effects due to age related changes in prefrontal connectively (MacPherson et al., 2002). Results of the current study and previous research are inconsistent with this hypothesis. As previously noted, however, verbal fluency performance requires verbal abilities, search and retrieval skills, adequate speed of processing, and an ability to inhibit inappropriate responses (Abwender et al., 2001; Henry & Crawford, 2004). Examining total word production on these tasks therefore may not provide a complete picture of fluency production in healthy aging. Consequently a process approach to interpreting verbal fluency results was adopted in this study using measures of clustering and switching.

Consistent with the study hypotheses, the older adults produced significantly fewer switches than the young and middle aged groups on the semantic task, and this effect approached significance on the phonemic task. There was no effect of age on average cluster size, whether single words were included or not. These results indicate that older adults compared to adults in the young and middle-aged groups demonstrate a lower ability to rapidly shift between clusters. Further, the variables proposed by Abwender and colleagues (2001) provided additional information in the current study. The older age group produced fewer hard switches on both fluency tasks but there was no age effect for cluster switches. Originally Abwender and colleagues (2001) proposed that hard switches are a reflection of the speeded nature of fluency measures, while cluster switches reflect strategic search processes. If this is true we would expect an age related decline in both hard switches and cluster switches due to age effects on processing speed and executive functioning, respectively (Henry & Phillips, 2006), a result not supported in this study. Alternatively, hard switches could be an indicator of processing speed and mental set shifting while cluster switching might be an indicator of intact semantic memory storage and intact connections among semantically related words. This interpretation is consistent with the results of the current study. The older age group showed a decline in hard switches, consistent with decreased processing speed and mental set shifting, but intact cluster switching, consistent with intact semantic memory storage and connectivity.

Taken together, the results of the current study lend support to the executive functioning hypothesis of healthy aging (MacPherson et al., 2002). Past research indicates that with increasing age, individuals show decreasing ability on measures of executive functioning, including on tasks of mental set shifting, as well as a decline in processing speed (Bryan &

Luszcz, 2000; Salthouse, 2010). In the current study, the older age group produced fewer total words due to a reduction in hard switching (a measure of processing speed and mental set shifting) with intact average cluster size scores and cluster switching (measures of memory storage and semantic/lexical connectivity).

In the current study, novel and repeated clusters showed age related effects for semantic but not for phonemic verbal fluency. Lanting and colleagues (2009) reported similar findings and concluded that returning to a previous cluster (i.e., repeated clusters) is a beneficial strategy for younger adults. This interpretation is plausible, however since adding the number of repeated clusters to the number of novel clusters would simply equal the number of total clusters produced, these variables do not appear to add additional information over and above the number of switches (i.e. the number of clusters generated minus one). Total switching, therefore, adequately captures this aspect of verbal fluency production in healthy aging.

The current study supports examining subcomponents of verbal fluency rather than just using a measure of total word production. While total word production does show age related change, at least for semantic fluency (Braaten et al., 2006; Bryan & Luszcz, 2000; Clark et al., 2009; Crossley, D'Arcy, & Rawson, 1997; Henry, Crawford, & Phillips, 2004), the reason for lower scores can be attributed to any number of cognitive or neuroanatomical differences (Abwender et al., 2001; Henry & Crawford, 2004). A process approach to the interpretation of verbal fluency performance, that examines specific subcomponents and strategies, has the potential to provide additional understanding of the cognitive abilities required to perform these tasks and how these abilities change with age. The results of the current study suggest that measures of clustering and switching provide additional information on age related changes in verbal fluency. In addition, these results indicate the importance of investigating verbal fluency variables with errors and perseverations (i.e., intrusions) excluded. In the current study, although there were no significant differences between age groups on number of intrusions, exclusion of these errors resulted in stronger effects in terms of predicted age group differences, supporting the assertion of Haugrud and colleagues (2010) that inclusion of errors and perseverations in the scoring of clustering and switching may artificially inflate the scores of older adults who tend to produce more intrusions than younger adults. Consequently, future research should examine verbal fluency variables excluding errors and persevations to fully understand age and clinical effects on clustering and switching subcomponents of these tasks.

A strength of the current study is that the three age groups were equivalent in terms of estimated verbal ability (as measured by the PPVT-R) and also were comparable in level of formal education. These findings are particularly relevant to phonemic fluency, which is highly sensitive to demographic differences (Crossley et al., 1997, Strauss et al., 2006). The findings from the current study demonstrate that phonemic fluency production is relatively age-insensitive in groups of adults equated on verbal intelligence and educational level.

Finally, the results of the current study support the use of a computer scoring program for the calculation of measures of clustering and switching to decrease the scoring time required for these procedures and to increase reliability. Our computer program reduced scoring time from approximately fifteen minutes per participant to one minute per participant, when compared to hand scoring. In addition, the use of a computer scoring program reduced the probability of coding errors, particularly since the scoring procedures for clustering and switching are complex. The use of a computer scoring program also allows for more detailed examination of a participant's verbal fluency production and additional variables can be created quickly based on the entered input. A detailed evaluation of this computer program is currently being prepared for publication and the protocol will be made available to increase ease and reliability of scoring both in research and clinical settings.

References

Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment, 8*, 323-336. doi: 10.1177/107319110100800308.

Benton, A. L., & Hamsher, K. (1989). *Multilingual aphasia examination.* Iowa City, Iowa: AJA Associates.

Brickman, A. M., Paul, R. H., Cohen, R. A., Williams, L. M., MacGregor, K. L., Jefferson, A., … Gordon, E. (2005). Category and letter verbal fluency across the adult lifespan: Relationship to EEG theta power. *Archives of Clinical Neuropsychology, 20*, 561-573. doi: 10.1016/j.acn.2004.12.006.

Braaten, A. J., Parsons, T. D., McCue, R., Sellers, A., & Burns, W. J. (2006). Neurocognitive differential diagnosis of dementing diseases: Alzheimer's dementia, vascular dementia, frontotemporal dementia, and major depressive disorder. *International Journal of Neuroscience, 116*, 1271-1293. doi: 10.1080/00207450600920928

Bryan, J., & Luszcz, M. A. (2000). Measurement of executive function: Consideration for detecting adult age differences. *Journal of Clinical and Experimental Neuropsychology, 22*, 40-55. doi: 10.1076/1380-3395%28200002%2922:1;1-8;FT040.

Clark, L. J., Gatz, M., Zheng, L., Chen, Y., McCleary, C., & Mack, W. J. (2009). Longitudinal verbal fluency in normal aging, preclinical, and prevalent Alzheimer's Disease. *American Journal of Alzheimer's Disease & Other Dementias, 24*, 461-468. doi: 10.1177/1533317509345154.

Crossley, M., D'Arcy, C., & Rawson, N. (1997). Letter and category fluency in community-dwelling Canadian seniors: A comparison of normal participants to those with dementia of the Alzheimer or vascular type. *Journal of Clinical and Experimental Neuropsychology, 19*, 52-62. doi: 10.1080/01688639708403836.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody picture vocabulary test revised.* Circle Pines, MN: American Guidance Service.

Haugrud, N., Lanting, S., & Crossley, M. (2010). The effects of age, sex and Alzheimer's disease on strategy use during verbal fluency tasks. *Aging, Neuropsychology, & Cognition, 17.* 220-239. doi: 10.1080/13825580903042700.

Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology, 18,* 284-295. doi: 10.1037/0894-4105.18.2.284

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia, 42,* 1212-1222. doi: 10.1016/j.neuropsychologia.2004.02.001

Henry, J. D., & Phillips, L. H. (2006). Covariates of production and perseveration on tests of phonemic, semantic and alternating fluency in normal aging. *Aging, Neuropsychology, and Cognition, 13,* 529-551. doi: 10.1080/138255890969537.

Hirshorn, E. A., & Thompson-Schill, S. L. (2006). Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency. *Neuropsychologia, 44,* 2547-2557. doi: 10.1016/j.neuropsychologia.2006.03.035.

Hughes, D. L., & Bryan, J. (2002). Adult age differences in strategy use during verbal fluency performance. *Journal of Clinical and Experimental Neuropsychology, 24,* 642-654. doi: 10.1076/jcen.24.5.642.1002.

Kavé, G. (2005). Phonemic fluency, semantic fluency, and difference scores: Normative data for adult Hebrew speakers. *Journal of Clinical and Experimental Neuropsychology, 27,* 690-699. doi: 10.1080/13803390490918499.

Lanting, S., Haugrud, N., & Crossley, M. (2009). The effects of age and sex on clustering and switching during speeded verbal fluency tasks. *Journal of the International Neuropsychological Society, 15,* 196-204. doi: 10.1017/S1355617709090237.

MacPherson, S.E., Phillips, L. H., & Della Sala, S. (2002). Age, executive function, and social decision making: A dorsolateral prefrontal theory of cognitive aging. *Psychology and Aging, 17,* 598-609. doi: 10.1037/0882-7974.17.4.598.

Mayer, U. (2002). On the dissociation between clustering and switching in verbal fluency: Comment on Troyer, Moscovitch, Winocur, Alexander and Stuss. *Neuropsychologia, 40,* 562-566. doi: 10.1016/S0028-3932%2801%2900132-4.

Mummery, C., Patterson, K., Hodges, J., & Wise, R. (1996). Generating "tiger" as an animal name or a word beginning with T: Differences in brain activation. *Proceedings: Biological Sciences, 263,* 989-995. Retrieved from http://www.jstor.org.cyber.usask.ca/stable/50587

Rodriguez-Aranda, C., & Martinussen, M. (2006). Age-related differences in performance of phonemic verbal fluency measured by Controlled Oral Word Association Task (COWAT): A meta-analytic study. *Developmental Neuropsychology, 30,* 697-717. doi: 10.1207/s15326942dn3002_3.

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society, 16,* 754-760. doi: 10.1017/S1355617710000706.

Struss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary, 3rd Ed.*. New York: Oxford University Press.

Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology, 22*, 370-378. doi: 10.1076/1380-3395%28200006%2922:3;1-V;FT370.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*, 138-146. doi: 10.1037/0894-4105.11.1.138.

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia, 36*, 499-504. doi: 10.1016/S0028-3932%2897%2900152-8.

Running head: CLUSTERING AND SWITCHING IN ALZHEIMER

Clustering and Switching Strategies During Verbal Fluency Performance Differentiate Alzheimer's Disease and Healthy Aging

Nicole Haugrud, Margaret Crossley, Mirna Vrbancic

This article has been published in the *Journal of the International Neuropsychological Society*, 2011, volume 17, page 1153-1157

Author Note

Nicole Haugrud, Department of Psychology, University of Saskatchewan; Margaret Crossley, Department of Psychology, University of Saskatchewan; Mirna Vrbancic, Department of Clinical Health Psychology, Royal University Hospital.

This research was supported in part by a Frederick Banting and Charles Best Canada Graduate Scholarship from the Canadian Institutes of Health Research awarded to Nicole Haugrud.

Correspondence concerning this article should be addressed to Nicole Haugrud, Department of Psychology, Arts 154, University of Saskatchewan, 9 Campus Drive, Saskatoon, Saskatchewan, Canada, S7N 5A5. Email: nicole.haugrud@usask.ca

Abstract

Clustering and switching strategies during phonemic and semantic verbal fluency tasks as defined by Troyer et al. (1997), Abwender et al. (2001) and Lanting et al. (2009) were compared using archival data to determine which scoring procedures best differentiate healthy older adults (n = 26) from individuals with early-stage Alzheimer's disease (AD, n = 26). Total word production showed the largest group difference, especially for semantic fluency. The AD group produced fewer switches when compared to the healthy control group, whereas the groups did not differ in cluster size. The AD group also accessed fewer novel semantic subcategories, presumably due to reduced access to semantic memory storage rather than lower processing speed. Clustering and switching scores on the phonemic task did not add information above total words produced, consistent with previous research indicating these variables are most informative in relation to semantic fluency.

Keywords: language, executive function, semantic memory, neuropsychology, dementia, cognitive

Clustering and Switching Strategies During Verbal Fluency Performance Differentiate Alzheimer's Disease and Healthy Aging

Alzheimer's disease (AD) typically presents with early impairments on tasks involving episodic memory and progresses to more global impairments including language and executive functioning (Braaten, Parsons, McCue, Sellers, & Burns, 2006). The most consistently found language deficit in early-stage AD is impaired word finding, particularly if given a target semantic category to guide the speeded generation of words (Braaten et al., 2006). Consequently, verbal fluency tests are frequently used in clinical settings to aid in the diagnosis of AD, and early-stage patients typically show greater semantic versus phonemic fluency impairment, presumably due to disproportionate effects in the temporal versus the prefrontal brain regions (Henry & Crawford, 2004).

Declines in semantic fluency total word production are found consistently in individuals with AD compared to healthy older adults, whereas the effect of AD on phonemic fluency performance is typically much smaller (Crossley, D'Arcy, & Rawson, 1997; Haugrud, Lanting, & Crossley, 2010: Henry, Crawford, & Phillips, 2004). In addition, individuals with AD produce fewer atypical or low frequency exemplars than normal adults (Sailor, Antoine, Diaz, Kuslansky, & Kluger, 2004).

Troyer, Moscovitch, and Winocur (1997) proposed that verbal fluency performance can be divided into clustering and switching components. Clustering involves the production of words within a semantic or phonemic subcategory and is proposed to rely primarily on temporal lobe processes. Switching refers to the ability to shift between clusters and is proposed to rely primarily on prefrontal lobe functions.

The model of Troyer et al. (1997) predicts that individuals with AD will show smaller cluster sizes with relatively intact switching rates due to decreases in efficient access to semantic knowledge. These results have been found by some researchers (Troyer et al., 1998), while other studies have only partially supported this theoretical difference (Haugrud et al., 2010). Previous researchers in this area have examined groups of individuals diagnosed with AD at varying stages of the disease (Beatty et al., 1997; Epker, Lacritz, & Munro Cullum, 1999; Haugrud et al., 2010; Troster et al., 1998; Troyer et al., 1998), which could explain differences between studies.

Modifications to the scoring procedures established by Troyer and colleagues (1997) have been proposed. For example, Abwender, Swan, Bowerman, and Connolly (2001) proposed

two types of switching strategies. Hard switching occurs between two single, non-clustered words or between a clustered word and a single word and is believed to result from the speeded nature of verbal fluency tasks. Cluster switching occurs between two groups of clustered words and is believed to reflect mental flexibility. Lanting and colleagues (2009) examined the number of novel clusters accessed, the number of clusters returned to in the same trial, and the percentage of clustered words. These variables were included to address limitations of the Troyer and colleagues (1997) model that included single words as a cluster with a score of zero. Finally, Haugrud and colleagues (2010) proposed that errors and perseverations should be removed from calculations of clustering and switching as these intrusions artificially inflate the cluster size scores for individuals with AD.

The current study used the methods of scoring proposed by Troyer et al. (1997), Abwender et al. (2001), and Lanting et al. (2009) to investigate verbal fluency in individuals diagnosed with early stage AD compared to healthy older adults. The current project had three goals: 1) to examine the variables of Abwender et al. (2001) and Lanting et al. (2009) in a group diagnosed with early-stage AD and, consistent with Haugrud et al. (2010), to examine these variables with errors removed; 2) to determine which of these scoring systems and variables best differentiate AD from healthy aging, contributing to our understanding of fluency decline in AD; and, 3) to use a computerized scoring procedure to generate clustering and switching variables in order to improve scoring accuracy and reliability.

Based on the two component model of verbal fluency and results from previous research, we hypothesized that the AD group would produce fewer total words on both verbal fluency tasks when compared to the healthy older adult group. Further, we hypothesized that the AD group would produce smaller average cluster sizes on both fluency tasks when compared to the healthy older adult group and fewer total switches on the semantic task. Due to disease-related effects on the semantic store (Braaten et al., 2006), we hypothesized that the AD group would produce fewer novel and repeated clusters, fewer cluster switches, and smaller percentage of clustered words than the healthy older adult group on the semantic task, but would show no differences on these variables on the phonemic task.

Methods

Participants

All data for this study were collected in compliance with the ethical regulations of the University of Saskatchewan. The current study used archival data from a subsample of participants (26 healthy older adults) recruited for a neuropsychological investigation of normal aging chosen for comparable age, years of education, and reading ability to 26 individuals diagnosed with AD according to the NINDS-ADRDA criteria (McKhann et al., 1984) recruited from an Aging Research and Memory Clinic. Results for total word production, average cluster size, and total switches, based on hand scoring of data from the current participants, have been reported previously by Haugrud and colleagues (2010). The current study extends this past work to include additional fluency variables not previously analyzed in an AD group. The healthy older adult group (15 females; 11 males) had a mean age of 70.5 (*SD* = 7.7) with an average of 11.9 (*SD* = 2.6) years of education. The Alzheimer's disease group (16 female; 10 males) had a mean age of 70.6 (*SD* = 7.6) with 11.4 years of education (*SD* = 3.4) and an average Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975) score of 24.7 (*SD* = 2.9). The groups did not differ in age, $F(1, 50) = 0.001$, $p = .971$, $\eta^2 = .001$, or education, $F(1, 50) = 0.410$, $p = .525$, $\eta^2 = .008$. Similarly, there was no significant difference between the healthy older adult group and the Alzheimer's disease group on the Wide Range Achievement Test-3 reading subtest (WRAT-3; Wilkinson, 1993), $F(1, 40) = 0.274$, $p = .604$, $\eta^2 = .007$, and the average scaled scores indicated average reading level for both groups ($M = 101.7$, $SD = 11.4$ and $M = 101.3$, $SD = 11.7$), for the normal and AD groups, respectively.

Materials

Participants completed the Controlled Oral Word Association Test (COWAT; Benton & Hamsher, 1989) as a measure of phonemic fluency and the Animal Naming test (AN; Spreen & Strauss, 1991) as a measure of semantic fluency.

Procedures and Scoring

The COWAT consists of three 60s trials requiring participants to produce as many words as possible that begin with the letters "C", "F", or "L." On the Animal Naming (AN) test, participants are given 60s to produce as many animal names as possible.

The verbal fluency variables were calculated both with intrusions (i.e., errors and perseverations) included and excluded. Detailed scoring procedures for the calculation of

average cluster size and number of switches have been previously reported (Troyer et al., 1997; 2000), as have the procedures for calculating hard and cluster switches (Abwender et al., 2001).

For the current study, a computer program was developed to generate the verbal fluency scores and to increase the reliability of the scoring procedures. The computer program is written in Python programming language and relies on word lists to group output according to scoring procedures. In a slight modification to the original scoring measures of Troyer and colleagues (1997), only the criterion of the same first two letters was used as a cluster for the phonemic task. The computer program was not able to score phonemic clusters that are homonyms, differ by a vowel sound or rhyme (Troyer et al., 1997). Using this computer scoring method, the verbal fluency scores were calculated quickly and were highly consistent with those obtained in previous studies using hand scoring methods, demonstrating the efficacy of the computer scoring program. Participant scores of average cluster size and number of switches differed slightly using the computer scoring program compared to the hand scoring method previously published in Haugrud et al. (2010). The largest difference was in the control group semantic average cluster size (M = 1.29, SD = 0.82 and M = 1.01, SD = .57, for Haugrud et al. [2010] and the current study, respectively) and the smallest was in the control group phonemic switches where the results were identical. The differences between the scores reported by Haugrud et al. (2010) and the current computer generated scores were not statistically significant and reflect slight modifications to the scoring procedures using the computer program and the challenges associated with reliably hand scoring these variables.

The calculations for the remaining variables have been described by Lanting and colleagues (2009). For the calculation of novel and repeated clusters, clusters were defined by the criteria of Troyer and colleagues (1997) for the semantic tasks, and included words with the same first two letters for the phonemic task. For the semantic task, when a word could be clustered into two different categories, the superordinate category of living environments was used, as described by Troyer and colleagues (1997). Novel and repeated clusters were calculated both including and excluding single words as clusters. Finally the percentage of clustered words per task was calculated.

Average cluster size was calculated according to the original method of Troyer et al. (1997) and re-calculated with single words excluded.

Results

For semantic and phonemic fluency tasks, separate one way analyses of variance (ANOVAs) were performed on all verbal fluency variables and partial η^2 is used as a measure of effect size. When errors and perseverations were removed from the calculation of verbal fluency variables, effect sizes for the significant findings were larger, and consistent with the hypotheses of the current study and past research. As a result, the following results are presented with intrusions removed.

Semantic fluency

Refer to Table 1 for the means and standard deviations of the semantic verbal fluency scores according to group. When compared to the healthy older adult group, the AD group produced significantly fewer total words, $F(1,50) = 42.854$, $p < .001$, and significantly fewer total switches, $F(1,48) = 24.831$, $p < .001$, hard switches, $F(1,50) = 10.244$, $p = .002$, and cluster switches, $F(1,50) = 7.050$, $p = .011$. The groups did not differ for semantic fluency average cluster size or for percentage of clustered words, but the AD group produced significantly fewer novel clusters, $F(1,50) = 20.154$, $p < .001$, and repeated clusters, $F(1,50) = 15.792$, $p < .001$, than the healthy older adult group, including fewer multiple word novel clusters, $F(1,50) = 16.583$, $p < .001$, and multiple word repeated clusters, $F(1,50) = 4.181$, $p = .046$. Examination of average cluster size excluding single words did not differentiate the AD group from the healthy older adult group.

Table 1

Semantic Verbal Fluency Scores (SD) for Participants with Alzheimer`s Disease (N=26) and for a Comparison Group of Healthy Older Adults (N=26)

	Healthy Controls	AD Group	partial η^2
Total Words Produced	18.5(4.9)	10.8(3.5)***	.462
Number of Switches	9.0(3.4)	5.1(2.2)***	.332
Average Cluster Size	1.01(0.57)	0.86(0.42)	.022
Average Cluster Size, no single words[a]	1.65(0.52)	1.47(0.54)	.031
Hard Switches	6.7(4.1)	3.8(2.2)**	.170
Cluster Switches	2.4(1.7)	1.3(1.1)*	.124
Novel Clusters	7.1(2.0)	5.0(1.4)***	.287
Repeated Clusters	2.9(2.0)	1.1(1.1)***	.240
Multiple Word Novel Clusters	4.4(1.6)	2.9(0.9)***	.249
Multiple Word Repeated Clusters	0.7(0.7)	0.3(0.6)*	.077
Percentage Clustered Words	72.6(18.5)	72.7(14.2)	.001
Number of Errors	0.3(0.6)	0.5(1.1)	.009
Number of Perseverations	1.1(2.0)	0.6(1.1)	.019

* $p < .05$, ** $p < .01$, ***$p < .001$

[a] average cluster size excluding single words

Phonemic fluency

Refer to Table 2 for the means and standard deviations of the phonemic verbal fluency scores according to group. The AD group produced fewer phonemic fluency total words than the healthy older adult group, $F(1,50) = 5.602$, $p = .022$. Groups did not differ on number of switches, number of hard or cluster switches, or on average cluster size. The AD group produced significantly fewer novel clusters, $F(1,50) = 4.992$, $p = .030$, and multiple word repeated clusters, $F(1,50) = 8.521$, $p = .005$, but there was no group difference for repeated clusters or on multiple word novel clusters. The AD group compared to the healthy older adult group produced significantly smaller average cluster size scores when single words were excluded, $F(1,50) = 8.878$, $p = .004$.

Table 2

Phonemic Verbal Fluency Variables for Participants with Alzheimer's Disease (N=26) and for a Comparison Group of Healthy Older Adults (N=26)

	Healthy Controls	AD Group	partial η^2
Total Words Produced	37.8(10.0)	29.9(13.7)*	.101
Number of Switches	23.8(7.8)	19.7(10.4)	.051
Average Cluster Size	0.48(0.22)	0.37(0.27)	.151
Average Cluster Size, no single words[a]	1.49(0.27)	1.13(0.56)**	.053
Hard Switches	22.2(7.8)	18.8(10.3)	.035
Cluster Switches	1.6(1.9)	0.8(1.2)	.059
Novel Clusters	13.6(2.1)	11.8(3.4)*	.091
Repeated Clusters	13.2(6.5)	10.8(7.6)	.028
Multiple Word Novel Clusters	5.6(2.1)	4.5(2.6)	.056
Multiple Word Repeated Clusters	1.7(1.4)	0.7(0.9)**	.146
Percentage Clustered Words	48.0(14.1)	39.8(19.7)	.057
Number of Errors	1.0(1.1)	0.9(1.2)	.003
Number of Perseverations	1.7(1.7)	2.2(2.5)	.012

* $p < .05$, ** $p < .01$, ***$p < .001$

[a] average cluster size excluding single word clusters

Discussion

Measures of effect size in the current study demonstrate that semantic fluency total word production best differentiates AD from healthy aging, closely followed by semantic fluency total switches. The variables of Abwender et al. (2001) did not add further information as both hard and cluster switching differentiated groups. Excluding single words from the analysis, consistent with Lanting et al. (2009), did not better differentiate AD from the healthy control group; however, the number of novel clusters accessed did differentiate the AD group from the healthy control group on both the phonemic and semantic tasks. Overall, clustering and switching variables showed larger effects in differentiating groups during semantic versus phonemic fluency tasks, indicating these variables are most informative when examining the effects of AD on semantic verbal fluency.

In contrast to the study hypothesis, healthy older adults and AD participants produced equivalent average cluster size scores during semantic fluency. Haugrud et al. (2010), using the same data set, found that males with AD, but not females, produced significantly smaller average cluster sizes than the healthy comparison group. Clarifying sex differences is an important direction for future research. Alternatively, contrasting findings might result from the use of the computerized scoring system in the current study that produced smaller differences in cluster size and switching scores compared to the hand scoring procedure used by Haugrud et al. (2010). Given that small changes in scoring consistency can change group effects on measures of clustering, average cluster size might not be the most effective method for differentiating AD from healthy aging. The current study was the first to use a computerized scoring system to calculate clustering and switching scores. Use of this program, in contrast to hand scoring procedures, was efficient and reliable and is strongly recommended for future research on clustering and switching variables.

McDowd and colleagues (2011) concluded that, compared to verbal ability, working memory, and inhibition, processing speed better predicts total correct responses and number of clusters produced on verbal fluency tasks in an older adult group. In the current study, lower fluency production in the AD group resulted from lower cluster production or switching rates. This effect was larger for novel clusters compared to repeated clusters. Novel cluster generation might therefore be a measure of intact semantic memory access in AD, rather than speed of processing. Alternatively, reduced switching rates for AD compared to normal participants could

result in the reduced number of novel clusters. Future research using regression modeling is needed to investigate measures of executive functioning, processing speed, and access to semantic memory as predictors of novel cluster generation in pathological and normal aging.

In summary, the current study found that total words, number of switches, and number of novel clusters best differentiate healthy older adults and AD participants, with the effects being larger on semantic compared to phonemic fluency. In addition, this study demonstrated the value of using a computerized scoring program to examine clustering and switching strategies in verbal fluency. Results should be replicated with a larger sample to support current findings and to investigate relationships among fluency variables and measures of processing speed, executive functions and semantic access for both normal and cognitively impaired males and females.

References

Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment, 8*, 323-336. doi: 10.1177/107319110100800308

Beatty, W. W., Testa, J. A., English, S., & Winn, P. (1997). Influences of clustering and switching on the verbal fluency performance of patients with Alzheimer's disease. *Aging, Neuropsychology, and Cognition, 4*, 273-279. doi: 10.1080/13825589708256652

Benton, A. L., & Hamsher, K. (1989). *Multilingual aphasia examination.* Iowa City, Iowa: AJA Associates.

Braaten, A. J., Parsons, T. D., McCue, R., Sellers, A., & Burns, W. J. (2006). Neurocognitive differential diagnosis of dementing diseases: Alzheimer's dementia, vascular dementia, frontotemporal dementia, and major depressive disorder. *International Journal of Neuroscience, 116*, 1271-1293. doi: 10.1080/00207450600920928

Crossley, M., D'Arcy, C., & Rawson, N. (1997). Letter and category fluency in community-dwelling Canadian seniors: A comparison of normal participants to those with dementia of the Alzheimer or vascular type. *Journal of Clinical and Experimental Neuropsychology, 19*, 52-62. doi: 10.1080/01688639708403836

Epker, M. O., Lacritz, L. H., & Munro Cullum, C. (1999). Comparative analysis of qualitative verbal fluency performance in normal elderly and demented populations. *Journal of Clinical and Experimental Neuropsychology, 21*, 425-434. doi: 10.1076/jcen.21.4.425.890

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-198. doi: 10.1016/0022-3956%2875%2990026-6

Haugrud, N., Lanting, S., & Crossley, M. (2010). The effects of age, sex and Alzheimer's disease on strategy use during verbal fluency tasks. *Aging, Neuropsychology, & Cognition, 17.* 220-239. doi: 10.1080/13825580903042700

Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology, 18*, 284-295. doi: 10.1037/0894-4105.18.2.284

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia, 42,* 1212-1222. doi: 10.1016/j.neuropsychologia.2004.02.001

Lanting, S., Haugrud, N., & Crossley, M. (2009). The effects of age and sex on clustering and switching during speeded verbal fluency tasks. *Journal of the International Neuropsychological Society, 15,* 196-204. doi: 10.1017/S1355617709090237

March, E. G., & Pattison, P. (2006). Semantic verbal fluency in Alzheimer's disease: Approaches beyond the traditional scoring system. *Journal of Clinical and Experimental Neuropsychology, 28*, 549-566. doi: 10.1080/13803390590949502

McDowd, J., Hoffman, L., Rozek, E., Lyons, K. E., Pahwa, R., Burns, J., & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychology, 25*, 210-225.

McKann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology, 34*, 939-944.

Sailor, K., Antoine, M., Diaz, M., Kuslansky, G., & Kluger, A. (2004). The effects of Alzheimer's disease on item output in verbal fluency tasks. *Neuropsychology, 18*, 306-314. doi: 10.1037/0894-4105.18.2.306

Spreen, O., & Strauss, E. (1991). *A compendium of neuropsychological tests*. New York: Oxford University Press.

Troster, A. I., Fields, J. A., Testa, J. A., Paul, R. H., Blanco, C. R., Hames, K. A.,...Beatty, W. W. (1998). Cortical and subcortical influences on clustering and switching in the performance of verbal fluency tasks. *Neuropsychologia, 36*, 295-304. doi: 10.1016/S0028-3932%2897%2900153-X

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*, 138-146. doi: 10.1037/0894-4105.11.1.138

Troyer, A. K., Moscovitch, M., Winocur, G., Leach, L., & Freedman, M. (1998). Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *Journal of the International Neuropsychological Society, 4*, 137-143. doi: 10.1017/S1355617798001374

Wilkinson, G. S. (1993). *Wide range achievement test – revision 3.* Wilmington, DE: Jastak Association.

Running head: LONGITUDINAL VERBAL FLUENCY IN ALZHEIMER'S DISEASE

Longitudinal Comparison of Verbal Fluency Subcomponents in Individuals Diagnosed with Alzheimer's Disease

Nicole Haugrud, Margaret Crossley, Mirna Vrbancic, Megan E. O'Connell, & Debra Morgan
University of Saskatchewan

Author Note

Nicole Haugrud, Department of Psychology, University of Saskatchewan; Margaret Crossley, Department of Psychology, University of Saskatchewan; Mirna Vrbancic, Department of Clinical Health Psychology, Royal University Hospital; Megan O'Connell, Department of Psychology, University of Saskatchewan; Debra Morgan, Canadian Centre for Health and Safety in Agriculture, University of Saskatchewan.

This research was supported in part by a Frederick Banting and Charles Best Canada Graduate Scholarship from the Canadian Institutes of Health Research awarded to Nicole Haugrud.

Correspondence concerning this article should be addressed to Nicole Haugrud, Department of Psychology, Arts 154, University of Saskatchewan, 9 Campus Drive, Saskatoon, Saskatchewan, Canada, S7N 5A5. Email: nicole.haugrud@usask.ca

Abstract

The aim of the current study was to compare subcomponents of verbal fluency production (i.e., clustering and switching variables) longitudinally in a group of individuals diagnosed with probable Alzheimer's disease (AD). Thirty-four individuals diagnosed with AD were assessed at initial diagnoses (Time 1) and at one-year follow up (Time 2). A subsample of 19 individuals was assessed at a two-year follow up assessment (Time 3). Participants completed Animal Naming (Strauss et al., 2006) as a measure of semantic fluency, and the Controlled Oral Word Association Test ("C", "F", and "L", Strauss et al., 2006) as a measure of phonemic fluency. Output was compared for the clustering and switching variables proposed by Troyer et al. (1997), Abwender et al. (2001), and Lanting et al. (2009). When all participants were included in the analysis, fluency results were inconsistent. Consequently participants were analyzed in two groups; those above clinical cut-off at initial assessment on a screening measure and those initially below cut-off. The group above initial clinical cut-off showed decline from Time 1 to Time 2 on phonemic fluency total words and an increase in phonemic fluency errors, with no change on the semantic task. For the subgroup followed over two years, those above initial cut-off showed decline on both phonemic and semantic fluency total words, and decline on phonemic switches, hard switches, and novel clusters. Participants initially below clinical cut-off showed more variability in performance, producing non-significant results on fluency variables. These results indicate that clustering and switching variables might be more useful in understanding early or preclinical decline in AD compared to later stages of the disease where all variables are impaired significantly and no longer show decline over time.

Keywords: longitudinal, verbal fluency, clustering, switching, Alzheimer's disease

Longitudinal Comparison of Verbal Fluency Subcomponents in Individuals Diagnosed with Alzheimer's Disease

The prevalence of dementia in Canada is expected to increase dramatically in the next 25 years (Alzheimer Society, 2010). Alzheimer's disease (AD) is the most common subtype of dementia (i.e., accounting for approximately 63% of cases), and over 480 000 Canadians current suffer from AD and related dementias (Alzheimer Society, 2010). Understanding the neuropsychological progression of AD is central to the development of strategies to help individuals with AD and their family members cope with the disease.

Individuals diagnosed with probable Alzheimer's disease (McKhann et al., 1984) show deficits compared to healthy older adults on a number of cognitive domains. For example, at the early stages of the disease, individuals with AD show severely impaired episodic memory, and mild to moderately impaired semantic memory and visuospatial skills (Braaten, Parsons, McCue, Sellers, & Burns, 2006; Hodges et al., 1999). Amnestic mild cognitive impairment (aMCI) is often considered a preclinical stage of AD and up to 80% of individuals diagnosed with aMCI convert to AD over a six year period (Petersen, 2004; Sarazin et al., 2007). Individuals who convert from aMCI to AD show poorer performance on measures of free and cued recall at initial assessment compared to individuals who do not convert to dementia (Mickes et al., 2007; Sarazin et al., 2007). As well, prior to diagnosis, when compared to individuals who do not later develop AD, older adults later diagnosed with AD show large effect size differences on measures of episodic memory, executive functioning, and processing speed, and small to medium effect size differences on measures of verbal ability, visuospatial skills, and attention (Backman, Jones, Berger, Laukka, & Small, 2005; Twamley, Ropacki, & Bondi, 2006). Longitudinal studies of AD progression after diagnosis reveal that individuals with more severe executive functioning deficits at initial diagnosis tend to show faster progression of decline (Musicco et al., 2010). Although with disease progression, impairment becomes more global. Together these studies indicate that prior to diagnosis individuals tend to show lower scores compared to healthy individuals on multiple cognitive domains, but these early impairments are most severe on episodic and semantic memory measures.

Degeneration of the medial temporal lobe structures, including the hippocampus and entorhinal cortex, has been associated with decline in memory and the learning of new information in AD patients (Braaten et al., 2006; Levy & Chelune, 2007; Rascovsky, Salmon,

Hansen, Thal, & Galasko, 2007; Scheff, Price, Schmitt, Scheff, & Mufson, 2011). In addition, imaging research has shown changes in the parietal lobe, frontal lobe, and posterior cingulate in individuals diagnosed with AD (Levy & Chelune, 2007; Twamley et al., 2006), even at preclinical or early stages of the disease, consistent with progression of AD to impaired attention, problem solving, mental set shifting, and visuospatial abilities.

Verbal fluency tasks are speeded word generation tasks frequently used in clinical settings to aid in the diagnosis of dementia. On these tasks participants are required to rapidly produce as many words as possible that either start with a specific letter such as “C”, “F”, or “L” (phonemic fluency) or belong to a specific semantic category such as animals (semantic fluency; Strauss, Sherman, & Spreen, 2006). Individuals with aMCI have shown impaired semantic and phonemic total word production compared to healthy older adults (Clark et al., 2009; Nutter-Upham et al., 2008; Raoux et al., 2008; Fagundo et al., 2008). However impaired fluency production is not always found in aMCI (Murphy, Rich, & Troyer, 2006). Research comparing individuals diagnosed with AD to healthy older adults consistently reveals impaired semantic verbal fluency performance (Crossley, D'Arcy, & Rawson, 1997; Haugrud, Lanting, & Crossley, 2010; Henry, Crawford, & Phillips, 2004; Laws, Duncan, & Gale, 2010; Mok, Lam, & Chiu, 2004). Although phonemic fluency performance also has been shown to decline in AD compared to healthy older adults, this effect typically is significantly smaller than for semantic verbal fluency (Canning, Leach, Stuss, Ngo, & Black., 2004; Clark et al., 2009; Crossley et al., 1997; Haugrud et al., 2010; Henry et al., 2004). In addition to lower total word generation, individuals with AD produce more errors on fluency tasks than healthy older adults (Marczinski & Kertesz, 2006), generate fewer atypical members of categories (Marczinski & Kertesz, 2006; Sailor, Antoine, Diaz, Kuslansky, & Kluger, 2004), and show a disrupted semantic network for animals (Chan, Salmon, & De La Pena, 2001; Hernandez, Costa, Juncadella, Sebastian-Galles, & Rene, 2008). Some longitudinal studies report impaired semantic fluency (Clark et al., 2009; Fagundo et al., 2008; Raoux et al., 2008) but intact phonemic fluency performance up to six years prior to AD diagnosis (Clark et al., 2009), while other researchers describe a decline in both fluency tasks in preclinical stages, but with a smaller effect for phonemic fluency (Mickes et al., 2007). Although there is longitudinal research examining verbal fluency in preclinical stages of AD, limited previous research has followed individuals longitudinally after diagnoses. One study found semantic fluency declines faster than phonemic fluency following AD diagnosis (Clark et

al., 2009). A second study found individuals diagnosed with AD produce more familiar words compared to healthy controls as the disease progresses (Moreno-Martinez & Montoro, 2010). Together, previous cross sectional research describes lower semantic and phonemic fluency in aMCI and AD compared to healthy older adults, with a smaller effect on the phonemic task, but longitudinal research on fluency tasks is limited, especially following AD diagnosis.

Although intact phonemic fluency performance is relatively more dependent on prefrontal lobe functioning and semantic fluency performance is relatively more dependent on medial temporal lobe functioning (Mummery et al., 1996), multiple cognitive components and associated brain regions are needed for normal performance on these tasks. For example, both tasks require verbal abilities, search and retrieval skills, adequate speed of processing, and an ability to inhibit inappropriate responses (Abwender, Swan, Bowerman, & Connolly, 2001; Henry & Phillips, 2006). An alternative approach to interpreting performance on neuropsychological assessment measures beyond a total score is a process approach that examines the subcomponents required for task performance. The two component model of verbal fluency production described by Troyer, Moscovitch, and Winocur (1997) is an example of this approach, and divides verbal fluency production into clustering and switching subcomponents. Clustering is the production of groups of semantically or phonemically related words (on the semantic and phonemic fluency test, respectively) and switching is the shifting between clusters of related words (Troyer et al., 1997). These authors propose that clustering is dependent on intact temporal lobe functioning while switching relies more heavily on prefrontal lobe functioning, a distinction supported by previous lesion and imaging research (Hirshorn & Thompson-Schill, 2006; Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998).

When examining groups of individuals diagnosed with Alzheimer's disease, some studies have reported impaired clustering on both phonemic and semantic fluency tasks, with preserved phonemic fluency switching (Haugrud et al., 2010; Troyer et al., 1998) or preserved semantic switching (March & Pattison, 2006), while other studies have reported both impaired clustering and switching on these tasks (Beatty, Testa, English, & Winn, 1997; Epker, Lacritz, & Cullum, 1999; Gomez & White, 2006; McDowd et al., 2011; Troster et al., 1998). Further, longitudinal comparisons with individuals diagnosed with aMCI or preclinical AD reveal impaired clustering and intact switching on semantic fluency tasks (Fagundo et al., 2008; Murphy et al., 2006), whereas Raoux et al. (2008) described individuals prior to AD diagnosis and reported the

opposite effect (i.e., intact clustering and impaired switching). Taken together, these previous results indicate a decline in both clustering and switching occurs in AD, particularly with disease progression. In early stage AD or aMCI results are less clear. In addition, other researchers have proposed examining additional fluency measures, including frequency of switching between single words and between clusters of more than one word (hard and cluster switches, respectively; Abwender et al., 2001) and the number of novel subcategories accessed during verbal fluency trials (Lanting, Haugrud, & Crossley, 2009). These additional variables have not been studied longitudinally in groups of early stage AD patients.

The goal of the current study was to analyze change in clustering and switching variables in an early stage AD group over a one or two year follow up period. Previous longitudinal studies of clustering and switching have compared individuals during the prediagnostic period (Fagundo et al., 2008; Raox et al., 2008). No previous research has followed AD individuals after diagnosis, or investigated the variables introduced by Abwender and colleagues (2001) and Lanting and colleagues (2009) longitudinally in an AD group. Consistent with previous research, it was hypothesized that individuals diagnosed with probable AD would show decline on phonemic and semantic fluency total words over the one and two year follow up periods, with a larger decline on semantic fluency. As well, consistent with increased disease effects on the medial temporal lobe, participants were hypothesized to show a decline in semantic fluency average cluster size over time, but little or no change in phonemic average cluster size. Semantic and phonemic fluency switching rates were hypothesized to show a progressive decline over time (i.e., from initial, to one- and two-year follow-up assessments) consistent with the progression of the disease in prefrontal lobe structures and associated connections. In addition, individuals with AD were hypothesized to show reduced hard switches, cluster switches, novel clusters, and repeated clusters with disease progression, with the largest effect evident during the two year follow up assessment.

Methods

Participants

Participants (N = 34; 28 females) were recruited from the Rural and Remote Memory Clinic in Saskatoon, Saskatchewan where they were referred for an initial interdisciplinary dementia assessment and followed for up to two years. Informed consent was obtained from patients and their caregivers for de-identified data to be incorporated into a larger database.

Participants were initially diagnosed by an interprofessional team with amnestic mild cognitive impairment (aMCI, *n* = 11) or probable Alzheimer's disease (*n* = 23) based on the recommendations for diagnostic criteria from the Third Canadian Consensus Conference on the Diagnosis and Treatment of Dementia (Rockwood, Bouchard, Camicioli, & Léger, 2007). To be included in the current study, all participants needed to be diagnosed with probable Alzheimer's disease by their one year follow up assessment. At initial assessment (Time 1) participants had an average age of 73.9 years (*SD* = 8.0) and an average of 10.9 (*SD* = 3.3) years of education. All participants completed a one year follow up assessment (Time 2) and a subsample of 19 participants completed a two year follow up assessment (Time 3). As part of a comprehensive neuropsychological assessment battery, participants completed the Modified Mini Mental State examination (3MS; Teng & Chui, 1987), the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, Tierney, Mohr, & Chase, 1998), the Stroop tasks (Strauss et al., 2006), the Trail Making Test Part A and B (Reitan, 1992), the Token test (Benton, Hamsher, & Sivan, 1994), the Controlled Oral Word Association Task (COWAT; Strauss et al., 2006) and the Animal Naming task (Animal Naming; Strauss et al., 2006). Table 1 shows the neuropsychological test battery raw data for the Time 1 and 2 data for the full sample of 34 participants, while Table 2 shows that neuropsychological test battery raw data for the subsample of 19 participants with Time 1, 2, and 3 data.

Table 1

Neuropsychological Assessment Battery Means (Standard Deviations) for Time 1 and Time 2 for the Full Participant Sample (N = 34) and Time 1, Time 2, and Time 3 Data for a Subsample of Participants Followed Over Two Years (n = 19)

Variable	Full Participant Sample		Subsample of Participants		
	Time 1	Time 2	Time 1	Time 2	Time 3
	(*N* = 36)	(*N* = 36)	(*n* = 19)	(*n* = 19)	(*n* = 19)
3MS[a]	80.2(8.4)	77.3(9.5)	81.8(9.6)	79.7(8.8)	76.9(11.5)
RBANS Immediate Memory Index[b]	69.7(14.2)	65.2(15.4)*	71.5(14.3)	65.7(15.7)	62.5(15.0)**
RBANS Visuospatial Index[b]	83.3(18.2)	88.2(19.1)	85.3(16.3)	91.5(17.6)	84.5(15.4)
RBANS Language Index[b]	87.4(13.7)	82.6(15.3)*	92.2(11.8)	86.0(15.0)	81.5(20.0)**
RBANS Attention Index[b]	81.6(17.3)	80.1(12.8)	81.1(14.3)	82.7(9.0)	81.8(13.0)
RBANS Delayed Memory Index[b]	52.3(9.1)	52.2(10.2)	52.7(9.0)	53.4(9.1)	50.1(11.0)
Stroop Color[c]	109.1(9.5)	111.3(2.1)	111.6(1.1)	111.6(0.6)	109.6(6.7)
Stroop Color Word[c]	58.1(21.6)	58.4(21.1)	65.5(17.3)	62.8(20.5)	56.3(24.7)
Trail Making Test Part A[d]	62.2(42.9)	57.9(28.0)	42.2(18.5)	55.2(31.8)	54.7(15.0)**
Trail Making Test Part B[d]	136.6(59.4)	154.4(70.2)	116.8(56.2)	147.6(56.2)	152.1(52.2)
Token Test[e]	41.1(3.3)	39.7(4.9)*	41.2(3.3)	39.3(6.2)	40.4(3.7)
COWAT[f]	27.7(11.8)	24.7(10.4)*	29.8(12.3)	27.1(9.8)	25.7(10.2)
Animal Naming[g]	11.2(4.5)	10.9(4.4)	12.9(5.1)	12.8(4.7)	11.5(4.8)

[a]The Modified Mini Mental State examination is scored out of a total of 100 (3MS; Teng & Chui, 1987). [b]The Repeatable Battery for the Assessment of Neuropsychological Status Index scores are scaled scores with a mean of 100 and a standard deviation of 10 (RBANS; Randolph et al., 1998). [c]The Stroop tasks are scored out of a total of 120 with the Color Task requiring individuals to read

color words printed in black ink while the Color-Word Task requires individuals to read color words printed in a discrepant color to the actual printed word (Strauss et al., 2006). [d]Scores for the Trail Making Test Part A are the number of seconds taken to sequentially join numbers in an array; scores for Part B are the number of seconds taken to alternate between joining numbers and letters in an array (Reitan, 1992). [e]The Token assesses comprehension of verbal commands that require individuals to respond by indicating specific shapes or colors of objects and has a total score of 45 (Benton, Hamsher, & Sivan, 1994). [f]Controled Oral Word Association Task requires production of words that begin with the letters "C", "F", or "L" in a 60 second trial; score is the number of correct words produced across three trials (COWAT; Strauss et al., 2006). [g]Animal Naming task requires production of names of animals in a 60 second trail; score is the number of correct words in one trial (Animal Naming; Strauss et al., 2006).

*$p < .05$, **$p < .01$ compared to Time 1 data

Materials

As part of the comprehensive neuropsychological research battery described above and in Table 1, participants completed the Controlled Oral Word Association Test (COWAT; Strauss et al., 2006) as a measure of phonemic fluency, and the Animal Naming test (AN; Strauss et al., 2006) as a measure of semantic fluency.

Procedures and Scoring

The COWAT and the Animal Naming test were administered according to standardized instructions (Strauss et al., 2006). On the COWAT, a measure of phonemic fluency, participants are required to produce as many words as possible that begin with the letters "C", "F", or "L". On the Animal Naming (AN) test, a measure of semantic fluency, participants are given 60s to produce as many animal names as possible.

Verbal fluency variables were calculated with intrusions (i.e., errors and perseverations) excluded consistent with previous research (Haugrud et al., 2010). The three trials were added together on the phonemic task to produce a phonemic total score for each variable with the exception of average cluster size where the three trials were averaged. Eight scores were generated for each fluency task (phonemic and semantic fluency); total words produced, total switches, average cluster size, hard switches, cluster switches, novel clusters, repeated clusters, and total intrusions. Total words produced was calculated as total words generated on a trial minus errors and perseverations (Strauss et al., 2006). Detailed scoring procedures for the calculation of clustering and switching variables have been previously reported (Abwender et al., 2001; Lanting et al., 2009; Troyer et al., 1997). Briefly, a cluster is a set of phonemically or semantically related words (on the phonemic or semantic task, respectively) while a switch is a shift between clusters. Hard switches (i.e., a switch between two single words or between a single word and clustered word), cluster switches (i.e., a switch between two groups of clustered words), novel clusters (i.e. the number of novel phonemic or semantic subcategories accessed, and repeated clusters (i.e. the number of repeated, previously accessed phonemic or semantic subcategories were also calculated. Total intrusions was calculated as the sum of all errors and repetitions on a trial. A computer program developed to calculate clustering and switching scores was used for the current study (see Study 1 of the current dissertation document).

Results

Paired samples t-tests were performed first for each verbal fluency variable comparing Time 1 to Time 2 for the full sample of participants. To determine whether a longer follow up period would produce more observed change paired samples t-tests were performed for each fluency variable comparing Time 1 to Time 3 for a subsample of participants who had Time 1, Time 2, and Time 3 data. Raw scores for the phonemic verbal fluency variables can be found in Table 2 and for the semantic verbal fluency variables in Table 3. Pearson's *r* is reported as a measure of effect size for the t-test results.

Table 2

Phonemic Verbal Fluency Means (Standard Deviations) For Time 1 and Time 2 For the Full Sample of Participants (N = 36) and For Time 1, Time 2, and Time 3 For a Subsample of Participants Followed Over Two Years (n = 19)

Variable	Full Participant Sample		Subsample of Participants		
	Time 1	Time 2	Time 1	Time 2	Time 3
	(*N* = 36)	(*N* = 36)	(*n* = 19)	(*n* = 19)	(*n* = 19)
Total Words Produced[a]	27.7(11.8)	24.7(10.4)*	29.8(12.3)	27.1(9.8)	25.7(10.2)
Total Switches[b]	17.7(9.0)	15.4(8.4)*	19.2(9.6)	17.8(8.2)	16.1(7.7)
Average Cluster Size[c]	0.44(0.35)	0.47(0.43)	0.49(0.44)	0.43(0.40)	0.38(0.24)
Hard Switches[d]	17.1(8.9)	14.7(8.1)*	18.5(9.7)	17.1(8.1)	15.2(7.1)
Cluster Switches[e]	0.7(0.8)	0.7(1.0)	0.6(0.8)	0.8(1.0)	0.9(1.2)
Novel Clusters[f]	11.8(3.1)	10.4(3.4)**	12.0(2.8)	11.5(3.3)	10.7(2.8)
Repeated Clusters[g]	8.9(6.7)	8.0(5.7)	10.2(7.5)	9.4(5.9)	8.4(6.0)
Total Intrusions[h]	1.8(2.0)	3.0(2.6)**	1.7(1.7)	2.3(1.9)	1.9(2.2)

[a]Total words produced is scored as the sum of all words on a 60 second trial minus errors and perseverations; the three phonemic trials are summed to produce a total phonemic score. [b]Total switches is scored as the number of clusters of words on a 60 second trial minus one; the three phonemic trials are summed to produce a total phonemic score. [c]Average cluster size is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one; the three phonemic trials are averaged to produce a total average phonemic score. [d]Hard switches is the sum of switches between two single words or between a single word and a clustered word on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [e]Cluster switches is scored as the sum of switches between two clusters of more than one word on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [f]Novel clusters is scored as the sum of new phonemic subcategories accessed

during a 60 second trial; the three phonemic trials are summed for a total phonemic score. [g]Repeated clusters is scored as the sum of all phonemic subcategories an individual returns to on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [h]Total intrusions is scored as the sum of all errors and perseverations across a trial; the three phonemic trials are summed to produce a total phonemic score

*$p < .05$; **$p < .01$

Table 3

Semantic Verbal Fluency Means (Standard Deviations) for Time 1 and Time 2 for the Full Participant Sample (N = 36) and for Time 1, Time 2, and Time 3 Data for a Subsample of Participants Followed over Two Years (n = 19)

Variable	Full Participant Sample		Subsample of Participants		
	Time 1	Time 2	Time 1	Time 2	Time 3
	(*N* = 36)	(*N*=36)	(*n*=19)	(*n*=19)	(*n*=19)
Total Words Produced[a]	11.2(4.5)	10.9(4.4)	12.9(5.1)	12.8(4.7)	11.5(4.8)
Total Switches[b]	5.3(2.3)	5.3(3.0)	6.0(2.6)	6.5(3.3)	6.0(3.1)
Average Cluster Size[c]	0.85(0.43)	0.85(0.44)	0.89(0.32)	0.81(0.41)	0.76(0.44)
Hard Switches[d]	4.1(2.3)	4.2(3.0)	4.5(2.3)	5.4(3.3)	4.9(3.3)
Cluster Switches[e]	1.2(1.0)	1.2(1.0)	1.5(1.1)	1.2(1.1)	1.1(1.4)
Novel Clusters[f]	4.9(1.6)	4.8(1.9)	5.4(1.9)	5.7(1.8)	5.1(2.0)
Repeated Clusters[g]	1.4(1.4)	1.5(1.6)	1.6(1.6)	1.8(1.9)	1.8(1.6)
Total Intrusions[h]	2.2(1.6)	1.3(1.5)*	2.3(1.5)	1.2(1.2)*	2.0(2.1)

[a]Total words produced is scored as the sum of all words on a 60 second trial minus errors and perseverations. [b]Total switches is scored as the number of clusters of words on a 60 second trial minus one. [c]Average cluster size is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one. [d]Hard switches is the sum of switches between two single words or between a single word and a clustered word on a 60 second trial. [e]Cluster switches is scored as the sum of switches between two clusters of more than one word on a 60 second trial. [f]Novel clusters is scored as the sum of new animal subcategories accessed during a 60 second trial. [g]Repeated clusters is scored as the sum of all animal subcategories an individual returns to on a 60 second trial. [h]Total intrusions is scored as the sum of all errors and perseverations across a trial

*$p < .05$

Phonemic fluency Time 1 versus Time 2

The decline from Time 1 to Time 2 on phonemic fluency total words produced was significant, $t(33) = 2.666$, $p = .012$, $r = .421$. Participants showed a significant increase in phonemic fluency errors from Time 1 to Time 2, $t(33) = -3.123$, $p = .004$, $r = .478$.

As with semantic total words produced, the number of words produced by the participants in the current study at Time 1 on phonemic fluency is significantly below published norms for healthy older adults (approximately one standard deviation below the average total word score for a healthy older adult group reported by Lanting et al., 2009).

There was signicicant decline from Time 1 to Time 2 on phonemic fluency total switches, $t(33) = 2.583$, $p = .014$, $r = .410$, hard switches, $t(33) = 2.567$, $p = .015$, $r = .408$, and novel clusters, $t(33) = 2.964$, $p = .006$, $r = .459$, but there was no effect for average cluster size, $t(33) = -0.342$, $p = .735$, $r = .059$, cluster switches, $t(33) = -0.150$, $p = .881$, $r = .026$, or repeated clusters, $t(33) = 1.467$, $p = .152$, $r = .247$.

Semantic fluency Time 1 versus Time 2

Participants showed no change in semantic fluency total word production from Time 1 to Time 2, $t(33) = 0.551$, $p = .585$, $r = .095$, or total intrusions, $t(33) = 2.405$, $p = .022$, $r = .386$. Although follow up over time indicates relative stability on semantic fluency, these scores at initial assessment are well below scores for a normal, healthy older adult group (the average score produced at Time 1 by participants in the current study for semantic total words produced is approximately two standard deviations below the mean for a healthy older adult group reported by Lanting et al., 2009). There was no observed decline from Time 1 to Time 2 in this sample, however the observed scores for the AD participants in this study are significantly impaired compared to healthy individuals.

There was no change in semantic fluency total switches, $t(33) = 0.076$, $p = .940$, $r = .013$, average cluster size, $t(33) = 0.026$, $p = .979$, $r = .005$, hard switches, $t(33) = -0.097$, $p = .924$, $r = .017$, cluster switches, $t(33) = 0.236$, $p = .815$, $r = .041$, novel clusters, $t(33) = 0.307$, $p = .761$, $r = .053$, or repeated clusters, $t(33) = -0.202$, $p = .841$, $r = .035$ from Time 1 to Time 2.

Phonemic fluency Time 1 versus Time 3 for a Subsample of Participants

For the subsample of participants with two year follow up data, the decline from Time 1 to Time 3 approached significance for phonemic fluency total words produced, $t(18) = 2.071$, $p = .051$, $r = .339$. There was no observed decline for phonemic fluency total intrusions, $t(18) = -$

0.380, $p = .709$, $r = .066$, total switches, $t(18) = 1.700$, $p = .106$, $r = .284$, average cluster size, $t(18) = 0.845$, $p = .409$, $r = .145$, hard switches, $t(18) = 1.827$, $p = .082$, $r = .303$, cluster switches, $t(18) = 0.793$, $p = .438$, $r = .137$, novel clusters, $t(18) = 1.955$, $p = .066$, $r = .322$, or repeated clusters, $t(18) = 1.264$, $p = .702$, $r = .215$.

Semantic fluency Time 1 versus Time 3 for a Subsample of Participants

No decline was observed from Time 1 to Time 3 for the subsample of participants with two year follow up data for semantic fluency total words produced, $t(18) = 1.603$, $p = .126$, $r = .269$, total intrusions, $t(18) = 0.661$, $p = .517$, $r = .144$, total switches, $t(18) = 0.074$, $p = .942$, $r = .013$, average cluster size, $t(18) = 1.033$, $p = .315$, $r = .177$, hard switches, $t(18) = 0.459$, $p = .652$, $r = .080$, cluster switches, $t(18) = 1.193$, $p = .249$, $r = .203$, novel clusters, $t(18) = 0.596$, $p = .559$, $r = .103$, or repeated clusters, $t(18) = 1.723$, $p = .102$, $r = .090$.

Verbal Fluency Comparisons Divided by Stage of Illness

In the current study, participant initial 3MS scores ranged from 56-99, indicating significant heterogeneity of stage of illness at initial assessment. Although all participants were recruited at initial assessment, participants showed variability in symptom severity at initial assessment. It is possible that this heterogeneity of initial symptom severity produced increased variability in test battery performance across individuals, eliminating any observable change over follow up. This is particularly likely given that, when all participants were included in the analysis, declines on the majority of neuropsychological test measures were non-significant, even over a longer two year follow up. To further investigate this possibility, participants were divided into two groups, one with 3MS scores greater than or equal to the clinical cut-off of 80 and one group with 3MS scores below this cut-off.

Comparison of fluency performance from Time 1 to Time 2 with participants divided by initial stage of illness.

Table 4 shows the means and standard deviations for Time 1 and Time 2 neuropsychological test battery scores for the full participant sample, divided by initial 3MS score above or below clinical cut-off. Table 5 shows the means and standard deviations for Time 1 and Time 2 phonemic verbal fluency variables for the full participant sample, divided by initial 3MS score above or below clinical cut-off. Table 6 shows the means and standard deviations for Time 1 and Time 2 semantic verbal fluency variables for the full participant sample, divided by initial 3MS score above or below clinical cut-off.

Table 4

Neuropsychological Test Battery Means (Standard Deviations) For Time 1 and Time 2 For the Full Participant Sample (N = 34) Divided by 3MS Score Into Above or Below Clinical Cut-off Groups

	Above Clinical Cut-off (*n* = 18)		Below Clinical Cutoff (*n* = 16)	
	Time 1	Time 2	Time 1	Time 2
3MS[a]	86.4(4.4)	80.6(7.6)*	73.1(5.7)	73.6(10.2)
RBANS Immediate Memory[b]	75.3(11.8)	68.3(15.6)**	64.6(13.8)	61.3(14.9)
RBANS Visuospatial[b]	93.7(14.0)	96.6(14.8)	71.6(14.8)	79.0(19.3)
RBANS Language[b]	92.2(11.2)	85.1(15.1)*	83.3(13.9)	79.4(15.4)
RBANS Attention[b]	88.4(16.8)	83.4(12.0)	74.9(13.8)	77.5(12.2)
RBANS Delayed Memory[b]	50.9(7.9)	49.0(10.4)	54.9(10.0)	55.9(8.9)
Stroop Color[c]	111.9(0.4)	111.0(2.7)	108.6(7.5)	111.7(0.7)
Stroop Color Word[c]	66.9(15.1)	57.5(21.9)	38.0(2.0)	55.7(13.6)
Trail Making Test Part A[d]	45.2(16.4)	48.9(19.4)	72.6(53.8)	70.2(35.4)
Trail Making Test Part B[d]	108.9(31.7)	141.7(52.2)**	184.7(69.0)	146.0(30.8)
Token Test[e]	42.3(2.7)	40.4(5.8)	39.6(3.0)	38.4(2.5)
COWAT[f]	32.8(10.5)	29.0(10.0)*	21.9(10.8)	19.8(8.6)
Animal Naming[g]	13.4(4.8)	12.8(4.7)	8.7(2.5)	8.8(3.1)

[a]The Modified Mini Mental State examination is scored out of a total of 100 (3MS; Teng & Chui, 1987). [b]The Repeatable Battery for the Assessment of Neuropsychological Status Index scores are scaled scores with a mean of 100 and a standard deviation of 10 (RBANS; Randolph et al., 1998). [c]The Stroop tasks are scored out of a total of 120 with the Color Task requiring individuals to read color words printed in black ink while the Color-Word Task requires individuals to read color words printed in a discrepant color to

the actual printed word (Strauss et al., 2006). [d]Scores for the Trail Making Test Part A are the number of seconds taken to sequentially join numbers in an array; scores for Part B are the number of seconds taken to alternate between joining numbers and letters in an array (Reitan, 1992). [e]The Token assesses comprehension of verbal commands that require individuals to respond by indicating specific shapes or colors of objects and has a total score of 45 (Benton, Hamsher, & Sivan, 1994). [f]Controled Oral Word Association Task requires production of words that begin with the letters "C", "F", or "L" in a 60 second trial; score is the number of correct words produced across three trials (COWAT; Strauss et al., 2006). [g]Animal Naming task requires production of names of animals in a 60 second trail; score is the number of correct words in one trial (Animal Naming; Strauss et al., 2006).

*$p < .05$; **$p < .01$

Table 5

Phonemic Verbal Fluency Means (Standard Deviations) for Time 1 and Time 2 for the Full Participant Sample (N = 34) Divided by 3MS Score Into Above or Below Clinical Cut-off Groups

	Above Clinical Cut-off (*n* = 18)		Below Clinical Cut-off (*n* = 16)	
	Time 1	Time 2	Time 1	Time 2
Total Words Produced[a]	32.8(10.5)	29.0(10.0)*	21.9(10.8)	19.8(8.6)
Total Switches[b]	20.9(8.2)	19.4(8.2)	14.1(8.8)	10.9(6.1)
Average Cluster Size[c]	0.47(0.34)	0.34(0.17)	0.41(0.38)	0.63(0.57)
Hard Switches[d]	20.3(8.3)	18.7(7.9)	13.4(8.5)	10.3(5.9)*
Cluster Switches[e]	0.6(0.7)	0.7(1.1)	0.7(0.9)	0.6(0.9)
Novel Cluster[f]	12.5(1.9)	11.7(3.3)	11.0(3.9)	9.0(3.1)*
Repeated Clusters[g]	11.4(7.1)	10.7(5.9)	6.1(5.1)	4.9(3.5)
Total Intrusions[h]	1.4(2.1)	3.3(3.2)**	2.3(1.8)	2.6(1.8)

[a]Total words produced is scored as the sum of all words on a 60 second trial minus errors and perseverations; the three phonemic trials are summed to produce a total phonemic score. [b]Total switches is scored as the number of clusters of words on a 60 second trial minus one; the three phonemic trials are summed to produce a total phonemic score. [c]Average cluster size is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one; the three phonemic trials are averaged to produce a total average phonemic score. [d]Hard switches is the sum of switches between two single words or between a single word and a clustered word on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [e]Cluster switches is scored as the sum of switches between two clusters of more than one word on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [f]Novel clusters is scored as the sum of new phonemic subcategories accessed during a 60 second trial; the three phonemic trials are summed for a total phonemic score. [g]Repeated clusters is scored as the sum of

all phonemic subcategories an individual returns to on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [h]Total intrusions is scored as the sum of all errors and perseverations across a trial; the three phonemic trials are summed to produce a total phonemic score

*$p < .05$; **$p < .01$

Table 6

Semantic Verbal Fluency Means (Standard Deviations) for Time 1 and Time 2 for the Full Participant Sample (N = 34) Divided by 3MS Score Into Above or Below Clinical Cut-off Groups

	Above Clinical Cut-off (*n* = 18)		Below Clinical Cut-off (*n* = 16)	
	Time 1	Time 2	Time 1	Time 2
Total Words Produced[a]	13.4(4.8)	12.8(4.7)	8.7(2.5)	8.8(3.1)
Total Switches[b]	6.4(2.3)	6.1(3.2)	4.1(1.7)	4.4(2.4)
Average Cluster Size[c]	0.85(0.31)	0.92(0.40)	0.85(0.55)	0.76(0.48)
Hard Switches[d]	4.9(2.4)	4.7(3.3)	3.3(1.8)	3.6(2.7)
Cluster Switches[e]	1.5(1.1)	1.5(1.0)	0.9(0.9)	0.8(0.9)
Novel Cluster[f]	5.6(1.7)	5.6(1.8)	4.2(1.2)	3.9(1.6)
Repeated Clusters[g]	1.8(1.6)	1.5(1.8)	0.9(0.9)	1.4(1.5)
Total Intrusions[h]	1.8(1.6)	1.2(1.4)	2.8(1.5)	1.4(1.7)*

[a]Total words produced is scored as the sum of all words on a 60 second trial minus errors and perseverations. [b]Total switches is scored as the number of clusters of words on a 60 second trial minus one. [c]Average cluster size is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one. [d]Hard switches is the sum of switches between two single words or between a single word and a clustered word on a 60 second trial. [e]Cluster switches is scored as the sum of switches between two clusters of more than one word on a 60 second trial. [f]Novel clusters is scored as the sum of new animal subcategories accessed during a 60 second trial. [g]Repeated clusters is scored as the sum of all animal subcategories an individual returns to on a 60 second trial. [h]Total intrusions is scored as the sum of all errors and perseverations across a trial

*$p < .05$

Phonemic verbal fluency performance from Time 1 to Time 2 for participants with an initial 3MS score above clinical cut-off.

Participants who scored above clinical cut off on the 3MS at initial assessment showed a decline from Time 1 to Time 2 on phonemic fluency total words produced, $t(17) = 2.540$, $p = .021$, $r = .524$, and an increase in the number of phonemic intrusions produced, $t(17) = -4.237$, $p = .001$, $r = .717$. There was no change from Time 1 to Time 2 for this group on phonemic fluency total switches, $t(17) = 1.443$, $p = .167$, $r = .330$, average cluster size, $t(17) = 1.449$, $p = .165$, $r = .332$, hard switches, $t(17) = 1.449$, $p = .165$, $r = .340$, cluster switches, $t(17) = 0.461$, $p = .651$, $r = .111$, novel clusters, $t(17) = 1.426$, $p = .172$, $r = .327$, or repeated clusters, $t(17) = 0.710$, $p = .487$, $r = .170$.

Semantic verbal fluency performance from Time 1 to Time 2 for participants with an initial 3MS score above clinical cut-off.

Participants with 3MS scores above clinical cut-off at initial assessment showed no decline from Time 1 to Time 2 on semantic fluency total words produced, $t(17) = 0.670$, $p = .512$, $r = .160$, total intrusions, $t(17) = 1.000$, $p = .331$, $r = .236$, total switches, $t(17) = 0.433$, $p = .670$, $r = .104$, average cluster size, $t(17) = 0.484$, $p = .634$, $r = .117$, hard switches $t(17) = 0.276$, $p = .786$, $r = .067$, cluster switches, $t(17) = 0.001$, $p = .999$, $r = .001$, novel cluster, $t(17) = 0.114$, $p = .911$, $r = .028$, or repeated clusters, $t(17) = 0.957$, $p = .352$, $r = .226$.

Phonemic verbal fluency performance from Time 1 to Time 2 for participants with an initial 3MS score below clinical cut-off.

Participants with an initial 3MS score below clinical cut-off showed a decline from Time 1 to Time 2 on number of total switches, $t(15) = 2.144$, $p = .049$, $r = .484$, and number of novel clusters, $t(15) = 2.739$, $p = .015$, $r = .577$. There was no change from Time 1 to Time 2 for phonemic fluency total words produced, $t(15) = 1.237$, $p = .235$, $r = .304$, total intrusions, $t(15) = 0.659$, $p = .520$, $r = .168$, average cluster size, $t(15) = -1.184$, $p = .255$, $r = .292$, hard switches, $t(15) = 2.076$, $p = .055$, $r = .472$, cluster switches, $t(15) = 0.194$, $p = .849$, $r = .050$, or repeated clusters, $t(15) = 1.409$, $p = .179$, $r = .342$.

Semantic verbal fluency performance from Time 1 to Time 2 for participants with an initial 3MS score below clinical cut-off.

Participants with an initial 3MS score below clinical cut-off showed no decline from Time 1 to Time 2 on semantic fluency total words produced, $t(15) = 0.115$, $p = .910$, $r = .030$,

total switches, $t(15) = 0.496$, $p = .627$, $r = .127$, average cluster size, $t(15) = 0.519$, $p = .611$, $r = .133$, hard switches, $t(15) = 0.594$, $p = .562$, $r = .152$, cluster switches, $t(15) = 0.436$, $p = .669$, $r = .112$, novel cluster, $t(15) = 0.436$, $p = .388$, $r = .224$, or repeated clusters, $t(15) = -1.142$, $p = .271$, $r = .283$. Participants produced significantly fewer errors at Time 2 compared to Time 1 on semantic fluency, $t(15) = 2.515$, $p = .024$, $r = .545$.

Comparison of fluency performance from Time 1 to Time 3 with participants divided by initial stage of illness.

Table 7 shows the neuropsychological test battery means and standard deviations for Time 1 and Time 3 for the subsample of participants with two year follow up data, divided by initial 3MS score above or below clinical cut-off. Table 8 shows the means and standard deviations for Time 1 and Time 3 phonemic verbal fluency variables for the subsample of participants with two year follow up data, divided by initial 3MS score above or below clinical cut-off. Table 9 shows the means and standard deviations for Time 1 and Time 3 semantic verbal fluency variables for the subsample of participants with two year follow up data, divided by initial 3MS score above or below clinical cut-off.

Table 7

Neuropsychological Test Battery Means (Standard Deviations) for Time 1 and Time 3 for the Subsample of Participants with Two Year Follow Up Data (n = 19) Divided by Initial 3MS Score Into Above or Below Clinical Cut-off Groups

	Above Clinical Cut-off (*n* = 12)		Below Clinical Cut-off (*n* = 9)	
	Time 1	Time 3	Time 1	Time 3
3MS[a]	87.5(5.0)	80.8(8.5)*	72.1(7.5)	70.3(13.4)
RBANS Immediate Memory[b]	76.2(12.1)	62.9(13.3)***	60.7(16.2)	61.7(19.4)
RBANS Visuospatial[b]	91.3(14.2)	89.2(9.7)	72.7(14.8)	75.2(21.2)
RBANS Language[b]	94.4(9.7)	84.1(16.8)**	81.8(18.2)	76.3(26.2)
RBANS Attention[b]	87.6(11.4)	83.3(11.6)	66.8(12.3)	78.2(16.8)
RBANS Delayed Memory[b]	53.0(8.2)	49.0(8.5)	47.3(8.9)	52.3(15.5)
Stroop Color[c]	111.7(0.9)	109.6(7.5)	110.0(2.8)	108.5(3.5)
Stroop Color Word[c]	72.4(11.8)	56.4(26.0)	NA	NA
Trail Making Test Part A[d]	38.5(11.0)	57.5(14.7)***	37.5(12.1)	54.5(10.5)**
Trail Making Test Part B[d]	102.4(25.6)	151.6(46.7)*	NA	NA
Token Test[e]	42.4(2.8)	42.0(3.0)	39.6(3.4)	36.8(2.4)
COWAT[f]	34.7(11.7)	27.1(10.5)**	21.4(8.4)	23.3(10.0)
Animal Naming[g]	14.8(5.2)	12.2(5.0)*	9.7(3.1)	10.3(4.5)

Note. NA: data were not available for Stroop Color Word, or Trail Making Test Part B for the subsample of participants with Time 3 data whose initial 3MS scores were below clinical cut-off.

[a]The Modified Mini Mental State examination is scored out of a total of 100 (3MS; Teng & Chui, 1987). [b]The Repeatable Battery for the Assessment of Neuropsychological Status Index scores are scaled scores with a mean of 100 and a standard deviation of 10

(RBANS; Randolph et al., 1998). [c]The Stroop tasks are scored out of a total of 120 with the Color Task requiring individuals to read color words printed in black ink while the Color-Word Task requires individuals to read color words printed in a discrepant color to the actual printed word (Strauss et al., 2006). [d]Scores for the Trail Making Test Part A are the number of seconds taken to sequentially join numbers in an array; scores for Part B are the number of seconds taken to alternate between joining numbers and letters in an array (Reitan, 1992). [e]The Token assesses comprehension of verbal commands that require individuals to respond by indicating specific shapes or colors of objects and has a total score of 45 (Benton, Hamsher, & Sivan, 1994). [f]Controled Oral Word Association Task requires production of words that begin with the letters "C", "F", or "L" in a 60 second trial; score is the number of correct words produced across three trials (COWAT; Strauss et al., 2006). [g]Animal Naming task requires production of names of animals in a 60 second trail; score is the number of correct words in one trial (Animal Naming; Strauss et al., 2006).

*$p < .05$; **$p < .01$; ***$p < .001$

Table 8

Phonemic Verbal Fluency Means (Standard Deviations) for Time 1 and Time 3 for the Subsample of Participants With Two Year Follow Up Data (n = 19) Divided by Initial 3MS Score Into Above or Below Clinical Cut-off Groups

	Above Clinical Cut-off (*n* = 12)		Below Clinical Cut-off (*n* = 7)	
	Time 1	Time 3	Time 1	Time 3
Total Words Produced[a]	34.7(11.7)	27.1(10.5)**	21.4(8.4)	23.3(10.0)
Total Switches[b]	22.4(9.0)	17.7(7.6)*	13.6(8.3)	13.4(7.7)
Average Cluster Size[c]	0.48(0.40)	0.32(0.14)	0.51(0.54)	0.48(0.34)
Hard Switches[d]	21.8(9.2)	16.7(6.9)*	12.9(8.2)	12.7(7.2)
Cluster Switches[e]	0.6(0.7)	1.0(1.2)	0.7(1.0)	0.7(1.1)
Novel Cluster[f]	12.8(2.0)	11.1(2.2)*	10.6(3.6)	10.1(3.8)
Repeated Clusters[g]	12.7(7.8)	9.6(6.6)	6.0(4.9)	6.3(4.5)
Total Intrusions[h]	0.8(10.0)	1.6(2.2)	3.3(1.5)	2.4(2.2)

[a]Total words produced is scored as the sum of all words on a 60 second trial minus errors and perseverations; the three phonemic trials are summed to produce a total phonemic score. [b]Total switches is scored as the number of clusters of words on a 60 second trial minus one; the three phonemic trials are summed to produce a total phonemic score. [c]Average cluster size is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one; the three phonemic trials are averaged to produce a total average phonemic score. [d]Hard switches is the sum of switches between two single words or between a single word and a clustered word on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [e]Cluster switches is scored as the sum of switches between two clusters of more than one word on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [f]Novel clusters is scored as the sum of new phonemic subcategories accessed during a 60 second trial; the three phonemic trials are summed for a total phonemic score. [g]Repeated clusters is scored as the sum of

all phonemic subcategories an individual returns to on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [h]Total intrusions is scored as the sum of all errors and perseverations across a trial; the three phonemic trials are summed to produce a total phonemic score

*$p < .05$; **$p < .01$

Table 9

Semantic Verbal Fluency Means (Standard Deviations) For Time 1 and Time 3 For the Subsample of Participants With Two Year Follow Up Data (n = 19) Divided by Initial 3MS Score Into Above or Below Clinical Cut-off Groups

	Above Clinical Cut-off (*n* = 12)		Below Clinical Cut-off (*n* = 7)	
	Time 1	Time 3	Time 1	Time 3
Total Words Produced[a]	14.8(5.2)	12.2(5.0)*	9.7(3.1)	10.3(4.5)
Total Switches[b]	6.9(2.5)	6.2(2.9)	4.4(2.0)	5.6(3.6)
Average Cluster Size[c]	0.89(0.31)	0.78(0.48)	0.89(0.35)	0.73(0.39)
Hard Switches[d]	5.2(2.6)	5.0(3.2)	3.4(1.5)	4.7(3.6)
Cluster Switches[e]	1.8(1.1)	1.2(1.6)	1.0(1.0)	0.9(1.1)
Novel Cluster[f]	5.9(2.0)	5.5(2.0)	4.6(1.4)	4.4(1.8)
Repeated Clusters[g]	2.0(1.8)	1.7(1.2)	0.9(0.9)	2.1(2.3)
Total Intrusions[h]	1.9(1.6)	1.1(1.4)	3.0(1.0)	3.6(2.1)

[a]Total words produced is scored as the sum of all words on a 60 second trial minus errors and perseverations. [b]Total switches is scored as the number of clusters of words on a 60 second trial minus one. [c]Average cluster size is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one. [d]Hard switches is the sum of switches between two single words or between a single word and a clustered word on a 60 second trial. [e]Cluster switches is scored as the sum of switches between two clusters of more than one word on a 60 second trial. [f]Novel clusters is scored as the sum of new animal subcategories accessed during a 60 second trial. [g]Repeated clusters is scored as the sum of all animal subcategories an individual returns to on a 60 second trial. [h]Total intrusions is scored as the sum of all errors and perseverations across a trial

*$p < .05$

Phonemic verbal fluency performance from Time 1 to Time 3 for participants with an initial 3MS score above clinical cut-off.

For the subsample of participants with Time 3 data with an initial 3MS score above clinical cut-off, there was significant decline observed from Time 1 to Time 3 for phonemic fluency total words produced, $t(11) = 4.288$, $p = .001$, $r = .791$, total switches, $t(11) = 2.344$, $p = .039$, $r = .577$, hard switches, $t(11) = 2.440$, $p = .033$, $r = .593$, and novel clusters, $t(11) = 3.079$, $p = .010$, $r = .680$. No change was observed from Time 1 to Time 3 on phonemic fluency average cluster size, $t(11) = 1.138$, $p = .279$, $r = .325$, cluster switches, $t(11) = 01.047$, $p = .318$, $r = .301$, repeated clusters, $t(11) = 1.636$, $p = .130$, $r = .442$, or total intrusions, $t(11) = -1.483$, $p = .166$, $r = .408$.

Semantic verbal fluency performance from Time 1 to Time 3 for participants with an initial 3MS score above clinical cut-off.

For the subsample of participants with Time 3 data with an initial 3MS score above clinical cut-off, there was a significant decline from Time 1 to Time 3 on semantic fluency total words produced, $t(11) = 2.340$, $p = .039$, $r = .576$. There was no observed decline on semantic fluency total switches, $t(11) = 0.974$, $p = .351$, $r = .282$, average cluster size, $t(11) = 0.689$, $p = .505$, $r = .203$, hard switches, $t(11) = 0.177$, $p = .863$, $r = .053$, cluster switches, $t(11) = 1.292$, $p = .223$, $r = .363$, novel clusters, $t(11) = 0.577$, $p = .576$, $r = .171$, repeated clusters, $t(11) = 0.670$, $p = .517$, $r = .198$, or total intrusions, $t(11) = 1.890$, $p = .085$, $r = .495$.

Phonemic verbal fluency performance from Time 1 to Time 3 for participants with an initial 3MS score below clinical cut-off.

For the subsample of participants with Time 3 data with an initial 3MS score below clinical cut-off there were no significant declines observed from Time 1 to Time 3 on phonemic fluency total words produced, $t(6) = 0.519$, $p = .622$, $r = .207$, total switches, $t(6) = 0.043$, $p = .967$, $r = .018$, average cluster size, $t(6) = 0.121$, $p = .908$, $r = .049$, hard switches, $t(6) = 0.045$, $p = .966$, $r = .018$, cluster switches, $t(6) = 0.001$, $p = .999$, $r = .001$, novel clusters, $t(6) = 0.300$, $p = .774$, $r = .122$, repeated clusters, $t(6) = 0.129$, $p = .902$, $r = .053$, or total intrusions, $t(6) = .779$, $p = .466$, $r = .303$.

Semantic verbal fluency performance from Time 1 to Time 3 for participants with an initial 3MS score below clinical cut-off.

For the subsample of participants with Time 3 data with an initial 3MS score below clinical cut-off there were no significant declines observed from Time 1 to Time 3 on semantic fluency total words produced, $t(6) = 0.464$, $p = .659$, $r = .203$, total switches, $t(6) = 0.834$, $p = .436$, $r = .322$, average cluster size, $t(6) = 0.749$, $p = .482$, $r = .292$, hard switches, $t(6) = 0.240$, $p = .818$, $r = .332$, cluster switches, $t(6) = 0.240$, $p = .818$, $r = .098$, novel clusters, $t(6) = 0.179$, $p = .864$, $r = .073$, repeated clusters, $t(6) = -1.264$, $p = .253$, $r = .459$, or total intrusions, $t(6) = 0.560$, $p = .596$, $r = .223$.

Discussion

The aim of the current study was to determine which subcomponents of verbal fluency production remain stable and which show decline over time in a group diagnosed with probable Alzheimer's disease (AD). Although there has been significant previous longitudinal research comparing individuals prior to AD diagnosis on verbal fluency tasks (Clark et al., 2009; Fagundo et al., 2008; Mickes et al., 2007; Raoux et al., 2008) the research following individuals after diagnosis has been more limited (Clark et al., 2009; Moreno-Martinez & Montoro, 2010). As well, only two previous studies have compared measures of average cluster size and number of switches longitudinally in individuals later diagnosed with AD (Fagundo et al., 2008; Raox et al., 2008) and no previous research has compared the variables of Abwender and colleagues (2001) or Lanting and colleagues (2009) in an AD group longitudinally.

Contrary to the study hypotheses, there were no significant changes from Time 1 to Time 2 in the full sample of AD participants or longitudinal decline from Time 1 to Time 3 in the subsample of AD participants tested over three years on the semantic fluency variables. The phonemic fluency task results were more consistent with the study hypotheses. From Time 1 to Time 2 participants showed decline on number of total words produced, total switches, hard switches, and novel clusters and produced more errors on phonemic fluency. For the subsample of participants followed over three years, participants showed decline from Time 1 to Time 3 on phonemic fluency total words and hard switches. However, participants in the current study showed a wide range in initial 3MS scores, indicating significant heterogeneity in disease severity at initial assessment. It is possible that the observed results comparing across all participants were confounded by large variability in initial disease severity.

To assess this hypothesis, participants were divided into two groups based on whether their initial 3MS scores were above or below clinical cut-off. Participants with initial 3MS scores

above clinical cut-off showed decline from Time 1 to Time 2 on phonemic fluency total words produced and an increase in number of phonemic errors over this same time period. No changes were observed on semantic fluency for this subgroup. For participants with two year follow up data who had initial 3MS scores above clinical cut-off, results showed decline from Time 1 to Time 3 on both phonemic and semantic total words produced, as well as phonemic fluency total switches, phonemic fluency hard switches, and phonemic fluency novel clusters. Participants whose initial 3MS scores were below clinical cut-off demonstrated more variability in performance both on the neuropsychological test battery and verbal fluency variables and analysis therefore showed minimal or no decline over time for this subgroup. These results indicate that the initial sample of 34 individuals represented a heterogeneous AD group. When participants were divided into two groups, one with initial 3MS scores above clinical cut-off (early stage group) and one with scores below clinical cut-off (later stage group), the early stage group showed more consistent decline in performance over time across both the neuropsychological test battery and verbal fluency scores. It is likely that those individuals at a later stage of disease at initial assessment showed larger variability in performance across individuals, masking any observed decline over time in this subsample. Alternatively, individuals at a later stage of disease severity at initial assessment might have already been experiencing significant impairment across the neuropsychological measures, and therefore further decline over only a one or two year follow up was not evident. Therefore results with respect to clustering and switching scores will only be interpreted for the early stage group.

It was hypothesized that declines in total word production would be evident on both fluency tasks over time and this decline would be greater on the semantic task. For the subsample of participants with initial 3MS scores above clinical cut-off, decline in phonemic fluency total words produced was observed both over a one and a two year follow up period while decline in semantic fluency total words produced was only observed over a longer, two year follow up. However, as noted in the results section, the AD group in the current study produced an average semantic fluency total score that was two standard deviations below published norms for a healthy older adult group, while the observed phonemic total score was only one standard deviation below published norms (Lanting et al., 2009). It is possible that semantic fluency production was already so severely impaired at initial assessment that participant scores were approaching floor effects and only further, slight decline was possible. In

contrast, phonemic verbal fluency total word production was relatively intact at initial assessment and therefore continued decline over time was observable. On the phonemic task, decreased total word production over follow up was the result of reduced switching (total switches and hard switches) and novel cluster production. Switching was proposed by Troyer and colleagues (1997) as a measure of speeded access to subcategories while hard switching was proposed to be a measure of processing speed by Abwender and colleagues (2001). Lanting and colleagues (2009) proposed novel clusters as a measure of search and retrieval ability. The observed decline on these measures in the current study is consistent with decline on measures of processing speed, memory search and retrieval, and mental set shifting with AD progression (Braaten et al., 2006). In contrast, there was no observed change in phonemic fluency average cluster size. Previous longitudinal research on these variables has focused on semantic fluency (Fagundo et al., 2008; Raoux et al., 2008) and has found conflicting results with respect to clustering and switching decline. The current study results are consistent with those of Raoux and colleagues (2009) that found decreased switching in preclinical AD but are at odds with the findings of Fagundo et al. (2009) that found decreased cluster size production.

The aim of the current study was to determine which clustering and switching variables show decline over time in a group diagnosed with Alzheimer's disease. For individuals at early stage of AD, phonemic fluency total switches, hard switches, and novel clusters appear to show decline over time producing a decline in phonemic total word production. Semantic fluency clustering and switching variables were severely impaired at initial assessment in the current study, even in a subgroup of individuals at early or pre-clinical stages of AD, and therefore clustering and switching variables on semantic fluency showed no further decline in this study. Clustering and switching are only two subcomponents of verbal fluency production that could be compared longitudinally in AD. For example previous research demonstrates that individuals with AD show reduced production of low frequency exemplars when compared to healthy individuals on semantic verbal fluency tasks (Marczinski & Kertesz, 2006; Sailor et al., 2004). Future research should compare frequency of exemplars produced by individuals diagnosed with AD longitudinally and how word frequency relates to clustering and switching scores.

In sum the results of the current study support a decline on phonemic fluency total word production due to declines in phonemic fluency switching with progressing of AD for individuals at early or preclinical stages of illness. In contrast, for individuals at later stages of

AD at initial diagnosis results were more variable and did not produce a consistent pattern of decline longitudinally. Future research should compare a larger sample of individuals at later stages of AD to determine whether phonemic fluency variables continue to decline.

Limitations of the current study is the short follow up period (maximum two years) and the small number of participants in the two year follow up group. As well, the heterogeneity of initial symptom severity contributed to the initial results with all participants included. Division of participants by initial stage of illness reduced this potential confound. However, the sample of participants followed over a two year period was further reduced in size by this division. Future research with a larger sample of individuals at later stages of the illness is warranted. A larger sample of participants would also allow for more detailed statistical analysis which would be more appropriate for longitudinal data analysis including growth model analysis.

References

Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment, 8*, 323-336. doi: 10.1177/107319110100800308

Alzheimer Society (2010). *Rising tide: The impact of dementia on Canadian society*. Retrieved from Alzheimer Society website: http://www.alzheimer.ca/english/rising_tide/rising_tide.htm

Backman, L., Jones, S., Berger, A. K., Laukka, E. L., & Small, B. J. (2005). Cognitive impairment in preclinical Alzheimer's disease: A meta-analysis. *Neuropsychology, 19*, 520-531. doi: 10.1037/0894-4105.19.4.520

Beatty, W. W., Testa, J. A., English, S., & Winn, P. (1997). Influences of clustering and switching on the verbal fluency performance of patients with Alzheimer's disease. *Aging, Neuropsychology, and Cognition, 4*, 273-279. doi: 10.1080/13825589708256652

Benton, A. L, Hamsher, K. de. S., & Sivan, A. B. (1994). *Multilingual Aphasia Examination (3rd ed.).* Iowa City, Iowa: AJA Associates.

Braaten, A. J., Parsons, T. D., McCue, R., Sellers, A., & Burns, W. J. (2006). Neurocognitive differential diagnosis of dementing diseases: Alzheimer's dementia, vascular dementia, frontotemporal dementia, and major depressive disorder. *International Journal of Neuroscience, 116*, 1271-1293. doi: 10.1080/00207450600920928

Canning, S. J., Leach, L., Stuss, D., Ngo, L., & Black, S. E. (2004). Diagnostic utility of abbreviated fluency measures in Alzheimer disease and vascular dementia. Neurology, 62, 556-562.

Chan, A. S., Salmon, D. P., & De La Pena, J. (2001). Abnormal semantic network for "animals" but not "tools" in patients with Alzheimer's disease. *Cortex, 37*, 197-217.

Clark, L. J., Gatz, M., Zheng, L., Chen, Y.L., McCleary, C., Mack, W.J. (2009). Longitudinal verbal fluency in normal aging, preclinical, and prevalent Alzheimer's disease. *American Journal of Alzheimer's Disease & Other Dementias, 24*, 461-468. doi: 10.1177/1533317509345154

Crossley, M., D'Arcy, C., & Rawson, N. S. B. (1997). Letter and category fluency in community-dwelling Canadian seniors: A comparison of normal participants to those

with dementia of the Alzheimer or vascular type. *Journal of Clinical and Experimental Neuropsychology, 19*, 52-62. doi: 10.1080/01688639708403836

Epker, M. O., Lacritz, L. H., & Munro Cullum, C. (1999). Comparative analysis of qualitative verbal fluency performance in normal elderly and demented populations. *Journal of Clinical and Experimental Neuropsychology, 21*, 425-434. doi: 10.1076/jcen.21.4.425.890

Fagundo, A. B., Lopez, S., Romero, M., Guarch, J., Marcos, T., & Salamero, M. (2008). Clustering and switching in semantic fluency: Predictors of the development of alzheimer's disease. *International Journal of Geriatric Psychiatry, 23*, 1007-1013. doi: 10.1002/gps.2025

Gomez, R. G., & White, D. A. (2006). Using verbal fluency to detect very mild dementia of the Alzheimer type. *Archives of Clinical Neuropsychology, 21*, 771-775. doi: 10.1016/j.acn.2006.06.012

Haugrud, N., Lanting, S., & Crossley, M. (2010). The effects of age, sex and Alzheimer's disease on strategy use during verbal fluency tasks. *Aging, Neuropsychology, & Cognition, 17.* 220-239. doi: 10.1080/13825580903042700.

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia, 42*, 1212-1222. doi: 10.1016/j.neuropsychologia.2004.02.001

Henry, J. D., & Phillips, L. H. (2006). Covariates of production and perseveration on tests of phonemic, semantic and alternating fluency in normal aging. *Aging, Neuropsychology, and Cognition, 13*, 529-551. doi: 10.1080/13825589-969537

Hernandez, M., Costa, A., Juncadella, M., Sebastian-Galles, N., & Rene, R. (2008). Category specific semantic deficits in Alzheimer's disease: A semantic priming study. *Neuropsychologia, 46*, 935-946.

Hirshorn, E. A., & Thompson Schill, S. L. (2006). Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency. *Neuropsychologia, 44*, 2547-2557. doi: 10.1016/j.neuropsychologia.2006.03.035

Hodges, J. R., Patterson, K., Ward, R., Garrard, P., Bak, T.,… Gregory, C. (1999). The differentiation of semantic dementia and frontal lobe dementia (temporal and frontal

variants of frontotemporal dementia) from early Alzheimer's disease: A comparative neuropsychological study. *Neuropsychology, 13,* 31-40. doi: 10.1037/0894-4105.13.1.31

Lanting, S., Haugrud, N., & Crossley, M. (2009). The effects of age and sex on clustering and switching during speeded verbal fluency tasks. *Journal of the International Neuropsychological Society, 15,* 196-204. doi: 10.1017/S1355617709090237

Laws, K. R., Duncan, A., & Gale, T. M. (2010). 'Normal' semantic-phonemic fluency discrepancy in Alzheimer's disease? A meta-analytic study. *Cortex, 46,* 595-601. doi: 10.1016/j.cortex.2009.04.009

Levy, J. A., & Chelune, G. J. (2007). Cognitive-behavioral profiles of neurodegenerative dementias: Beyond Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology, 20,* 227-236. doi: 10.1177/0891988707309906

March, E. G., & Pattison, P. (2006). Semantic verbal fluency in Alzheimer's disease: Approaches beyond the traditional scoring system. *Journal of Clinical and Experimental Neuropsychology, 28,* 549-566. doi: 10.1080/13803390590949502

Marczinski, C. A., & Kertesz, A. (2006). Category and letter fluency in semantic dementia, primary progressive aphasia, and Alzheimer's disease. *Brain and Language, 97,* 258-265. doi: 10.1016/j.bandl.2005.11.001

McDowd, J., Hoffman, L., Rozek, E., Lyons, K. E., Pahwa, R., Burns, J., & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychology, 25,* 210-225. doi: 10.1037/a0021531

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology, 34,* 939-944.

Mickes, L., Wixted, J. T., Fennema-Notestine, C., Galasko, D., Bondi, M. W., Thal, L. J., & Salmon, D. P. (2007). Progressive impairment on neuropsychological tasks in a longitudinal study of preclinical Alzheimer's disease. *Neuropsychology, 21,* 696-705. doi: 10.1037/0894-4105.21.6.696

Mok, E. H. L., Lam, L. C. W., & Chiu, H. F. K. (2004). Category verbal fluency test performance in Chinese elderly with Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 18,* 120-124. doi: 10.1159/000079190

Moreno-Martinex, F. J., & Montoro, P. R. (2010). Longitudinal patterns of fluency impairment in dementia: The role of domain and "nuisance variables". *Aphasiology, 24,* 1389-1399. doi: 10.1080/02687030903515370

Murphy, K. J., Rich, J. B., & Troyer, A. K. (2006). Verbal fluency patterns in amnestic mild cognitive impairment are characteristic of Alzheimer's type dementia. *Journal of the International Neuropsychological Society, 12,* 570-574. doi: 10.1017/S1355617706060590

Musicco, M., Salamone, G., Caltagirone, C., Cravello, L., Fadda, L., Lupo, F., Mosti, S., Perri, R., & Palmer, K. (2010). Neuropsychological predictors of rapidly progressing patients with Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 30,* 219-228. doi: 10.1159/000319533

Nutter-Upham, K. E., Saykin, A. J., Rabin, L. A., Roth, R. M., Wishart, H. A., Pare, N., & Flashman, L. A. (2008). Verbal fluency performance in amnestic MCI and older adults with cognitive complaints. *Archives of Clinical Neuropsychology, 23,* 229-241. doi: 10.1016/j.acn.2008.01.005

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine, 256,* 183-194.

Price, C. C., Jefferson, A. L., Merino, J. G., Heilman, K. M., & Libon, D. J. (2005). Subcortical vascular dementia: Integrating neuropsychology and neuroradiologic data. *Neurology, 65,* 376-382. doi: 10.1212/01.wnl.0000168877.06011.15

Randolph, C., Tierney, M. C., Mohr, E., & Chase, T. N. (1998). The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): Preliminary clinical validity. *Journal of Clinical and Experimental Neuropsychology, 20,* 310-319.

Raoux, N., Amieva, H., Le Goff, M., Auriacombe, S., Carcaillon, L.,… Dartigues, J. F. (2008). Clustering and switching processes in semantic verbal fluency in the course of Alzheimer's disease subjects: Results from the PAQUID longitudinal study. *Cortex, 44,* 1188-1196. doi: 10.1016/j.cortex.2007.08.019

Rascovsky, K., Salmon, D. P., Hansen, L. A., Thal, L. J., & Galasko, D. (2007). Disparate letter and semantic category fluency deficits in autopsy-confirmed frontotemporal dementia and Alzheimer's disease. *Neuropsychology, 21,* 20-30. doi: 10.1037/0894-4105.21.1.20

Reitan, R. M. (1992). *Trail Making Test: manual for administration and scoring*. Arizona: Reitan Neuropsychology Laboratory.

Rockwood, K., Bouchard, R. W., Camicioli, R., & Leger, G. (2007). Toward a revision of criteria for the dementias. *Alzheimer's & Dementia, 3*, 428-440. doi: 10.1016/j.jalz.2007.07.014

Sailor, K., Antoine, M., Diaz, M., Kuslansky, G., & Kluger, A. (2004). The effects of Alzheimer's disease on item output in verbal fluency tasks. *Neuropsychology, 18*, 306-314. doi: 10.1037/0894-4105.18.2.306

Sarazin, M., Berr, C., De Rotrou, J., Fabrigoule, C., Pasquier, F., Legrain, S.,…Dubous, B. (2007). Amnestic syndrome of the medial temporal type identifies prodromal AD: A longitudinal study. *Neurology, 69*, 1859-1867.

Scheff, S. W., Price, D. A., Schmitt, F. A., Scheff, M. A., & Mufson, E. J. (2011). Synaptic loss in the inferior temporal gyrus in mild cognitive impairment and Alzheimer's disease. Journal of Alzheimer's Disease, 24, 547-557. Doi: 10.3233/JAD-2011-101782

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary, 3rd Ed.*. New York: Oxford University Press.

Teng, E. L., & Chui, H. C. (1987). The modified Mini-Mental State (3MS) Examination. *Journal of Clinical Psychiatry, 48*, 314-318.

Troster, A. I., Fields, J. A., Testa, J. A., Paul, R. H., Blanco, C. R., Hames, K. A.,...Beatty, W. W. (1998). Cortical and subcortical influences on clustering and switching in the performance of verbal fluency tasks. *Neuropsychologia, 36*, 295-304. doi: 10.1016/S0028-3932%2897%2900153-X

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*, 138-146. doi: 10.1037/0894-4105.11.1.138

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia, 36*, 499-504.

Twamley, E. W., Ropacki, R. A. L., & Bondi, M. W. (2006). Neuropsychological and neuroimaging changes in preclinical Alzheimer's disease. *Journal of the International Neuropsychological Society, 12*, 707-735. doi: 10.1017/S135561770606863

Running head: PATTERNS OF VERBAL FLUENCY PRODUCTION

Patterns of Verbal Fluency Production Differentiate Subtypes of Dementia and Healthy Aging

Nicole Haugrud, Margaret Crossley, Mirna Vrbancic, Megan E. O'Connell, & Debra Morgan
University of Saskatchewan

Author Note

Nicole Haugrud, Department of Psychology, University of Saskatchewan; Margaret Crossley, Department of Psychology, University of Saskatchewan; Mirna Vrbancic, Department of Clinical Health Psychology, Royal University Hospital; Megan E. O'Connell, Department of Psychology, University of Saskatchewan; Debra Morgan, Canadian Centre for Health and Safety in Agriculture, University of Saskatchewan.

This research was supported in part by a Frederick Banting and Charles Best Canada Graduate Scholarship from the Canadian Institutes of Health Research awarded to Nicole Haugrud.

Correspondence concerning this article should be addressed to Nicole Haugrud, Department of Psychology, Arts 154, University of Saskatchewan, 9 Campus Drive, Saskatoon, Saskatchewan, Canada, S7N 5A5. Email: nicole.haugrud@usask.ca

Abstract

The effects of dementia on measures of phonemic (Controlled Oral Word Association Test) and semantic (Animal Naming) verbal fluency production were compared for clustering and switching scores as defined by Troyer et al. (1997), Abwender et al. (2001) and Lanting et al (2009). Healthy older adults (n = 26) were compared to patients diagnosed with amnestic mild cognitive impairment (aMCI; n = 14), Alzheimer's disease (AD; n = 22), vascular dementia (VaD; n = 23), Lewy Body dementia (DLB; n = 11), behavioural-variant frontotemporal dementia (FTD-bv; n = 10), and language-variant frontotemporal dementia (FTD-lang; n = 10). The aMCI group performed normally on all fluency measures except semantic total word production. All dementia groups had impaired total word production on semantic fluency compared to healthy older adults. The AD group also produced smaller cluster sizes and fewer switches on semantic fluency. The VaD and DLB groups were impaired on all measures except cluster size. The FTD-bv group was more impaired on phonemic than semantic fluency, and especially during phonemic switching. The FTD-lang group showed consistent impairment across all measures and produced the largest number of errors. Total word production was a sensitive but not specific measure of dementia in this study, whereas clustering and switching strategies differentiated dementia subtypes. Results are consistent with impaired semantic memory storage in AD, impaired processing speed and set shifting in DLB and VaD, and impaired complex strategic search processes in FTD-bv.

Keywords: dementia, verbal fluency, clustering, switching, Alzheimer's disease

Patterns of Verbal Fluency Production Differentiate Subtypes of Dementia and Healthy Aging

With the anticipated increase in the prevalence of Alzheimer's disease (AD) and related dementias as the "baby boomers" age into older adulthood (Alzheimer Society, 2010) there is an increased need for accurate differential diagnosis of dementia subtypes. Individuals with diagnoses of different dementia sub-types demonstrate unique behavioural, cognitive, and functional impairments (e.g., Robillard, 2007). Accurate diagnosis at early stages of the disease offers the best hope for effective treatment and management strategies. Neuropsychological assessments provide important information for the in vivo diagnosis of dementia. There is significant overlap in the clinical presentations of dementia subtypes, however, making reliable diagnosis difficult, particularly at early stages of decline (Braaten, Parsons, McCue, Sellers, & Burns, 2006). One neuropsychological assessment measure that shows this ambiguity in differentiating types of dementia is verbal fluency. Verbal fluency tests are speeded word generation tasks requiring participants to generate as many words as possible either starting with a certain letter (phonemic fluency) or belonging to a specific semantic category (semantic fluency). These tests are sensitive measures for detecting cognitive impairment due to dementia (Braaten et al., 2006; Levy & Chelune, 2007). In addition to total word production, verbal fluency tasks can be evaluated by comparing patterns, strategies, or subcomponents of word generation. The aim of the current study was to determine which subcomponents of verbal fluency production have utility in differentiating dementia sub-types.

Amnestic mild cognitive impairment (aMCI) is characterized by subjective memory complaints and poor performance on measures of episodic memory in the absence of global cognitive decline or significant problems with activities of daily living (Petersen, 2004). Reports of verbal fluency performance in groups diagnosed with aMCI have been inconsistent. Some studies report impaired semantic (Nutter-Upham et al., 2008; Raoux et al., 2008; Fagundo et al., 2008) and phonemic (Nutter-Upham et al., 2008) total word production in aMCI groups compared to healthy older adults, while other studies have failed to show phonemic or semantic total word decline in aMCI (Murphy, Rich, & Troyer, 2006). aMCI is often considered a preclinical stage of Alzheimer's disease (AD) because up to 80% of individuals diagnosed with aMCI convert to AD over a six year period (Petersen, 2004; Sarazin et al., 2007). Individuals with AD show impairment on measures of episodic memory and confrontational naming at early

stages of the disease, and progress to more global impairments at the later stages (Braaten et al., 2006; Giovagnoli, Erbetta, Reati, & Bugiani, 2008). Declines in semantic verbal fluency performance have been found in Alzheimer's disease (AD) patients compared to healthy older adults (Crossley, D'Arcy, & Rawson, 1997; Haugrud, Lanting, & Crossley, 2010; Henry, Crawford, & Phillips, 2004; Laws, Duncan, & Gale, 2010; Mok, Lam, & Chiu, 2004). Phonemic fluency performance also has been shown to decline in AD compared to healthy older adults, but this effect is smaller when compared to the semantic task (Canning, Leach, Stuss, Ngo, & Black., 2004; Crossley et al., 1997; Haugrud et al., 2010; Henry et al., 2004). Studies with less severely impaired AD groups compared to studies with more advanced or mixed groups of patients tend to show better phonemic fluency performance (Laws et al., 2010). In addition to lower total word generation, individuals with AD produce more errors on fluency tasks than healthy control groups (Marczinski & Kertesz, 2006).

The term Vascular Cognitive Impairment (VCI) is becoming increasingly accepted as a broader term encompassing all forms of cognitive loss due to cerebrovascular disease (Rockwood, Bouchard, Camicioli, & Leger, 2007). VCI-no dementia, subcortical vascular dementia (VaD) with white matter changes on neuroimaging, and VaD with multiple or single infarcts are three recognized subtypes of VCI (Rockwood et al., 2007). The National Institute of Neurologic Disorders and Stroke and the Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN) criteria for VaD require a diagnosis of dementia (including decline from a previous level of functioning and impairment on memory and two or more cognitive domains), evidence based on neuroimaging of cerebrovascular disease, and a convincing relationship between dementia presentation and the progression of cerebrovascular disease (Roman et al., 1993). However, the Third Canadian Consensus Conference on the Diagnosis and Treatment of Dementia (CCCDTDT) reported these criteria show high specificity at a cost of low sensitivity (Robillard, 2007). Other criteria for VaD have been proposed, however similar criticisms to the NINDS-AIREN criteria have been reported (Robillard, 2007) and there is no current consensus on the preferred system. Given the multiple potential neuroanatomical causes of VCI, a consistent neuropsycholoical profile for this disease is unrealistic. Nevertheless, previous research supports a dysexecutive profile for VaD (Rockwood et al., 2007), as well as slowed speed of processing, lowered sustained attention, cognitive inflexibility, and relatively intact episodic memory (Lafosse et al., 1997; Levy &

Chelune, 2007; Robillard, 2007). Groups of VaD patients show decreased word production on both phonemic and semantic fluency tasks compared to healthy older adults (Braaten et al., 2006). As a result of this equivalent decline on both tasks in VaD, individuals with VaD tend to have lower output than individuals with AD on phonemic fluency tasks (Canning et al., 2004; Lafosse et al., 1997; Levy & Chelune, 2007).

Another prominent dementia subtype is dementia with Lewy bodies (DLB). The core features of DLB are fluctuating cognition with variations in attention and alertness, recurrent visual hallucinations that are well formed, and spontaneous motor features of parkinsonism (Robillard, 2007). On neuropsychological measures, individuals with DLB show prominent deficits in visuospatial ability, attention, and executive functioning (Levy & Chelune, 2007; Oda, Yamamoto, & Maeda, 2009; Troster, 2008). Individuals with DLB also tend to show impairment on both semantic and phonemic fluency and are more impaired than individuals with AD on phonemic tasks (Levy & Chelune, 2007; Ralph et al., 2000).

Frontotemporal dementia (FTD) can be divided into three clinical subtypes. Behavioral variant FTD (FTD-bv) is characterized by prominent changes in social behaviour and personality as a consequence of orbitobasal prefrontal lobe degeneration (Robillard, 2007). Two language variants of FTD are semantic dementia (SD) and progressive nonfluent aphasia (FTD-pnf). Individuals with SD show fluent but empty spontaneous speech and a breakdown in language comprehension due to left anterolateral temporal lobe atrophy (Robillard, 2007). FTD-pnf is characterized by impaired phonologic and syntactic language components due to left perisylvian atrophy (Robillard, 2007). Decreased word production on both semantic and phonemic fluency tasks has been found in individuals with frontotemporal dementia, with more severe impairment on the phonemic task (Hodges et al., 1999; Levy & Gordon, 2007; Rascovsky, Salmon, Hansen, Thal, & Galasko, 2007). Within subgroups of FTD, individuals with FTD-pnf tend to produce the fewest words on verbal fluency tasks, followed next by FTD-SD and then by individuals with FTD-bv (Marczinski & Kertesz, 2006).

In summary, individuals diagnosed with AD tend to show more severely impaired semantic fluency compared to phonemic fluency, while individuals with FTD tend to show the opposite pattern (e.g Henry et al., 2004; Hodges et al., 1999; Levy & Gordon, 2007). Results for aMCI are mixed, with some studies showing phonemic or semantic total word impairment (Raoux et al., 2008; Fagundo et al., 2008) while other studies show no difference when

compared to healthy older adult groups (Murphy et al., 2006). For individuals with VaD and DLB, both fluency tasks appear equally impaired (Levy & Chelune, 2007; Ralph et al., 2000). In conclusion, total word production on verbal fluency tasks can be helpful in differentiating AD from FTD, but these total output measures are less effective in identifying VaD, DLB or aMCI.

An alternative approach to the interpretation of neuropsychological assessment data beyond simply comparing total scores on measures is to use a process approach to interpretation. This approach examines the components of a task required for normal performance. Through this method specific strategies and approaches to a task can be compared to provide additional information over and above group differences on total scores. The two component model of verbal fluency production described by Troyer, Moscovitch, and Winocur (1997) is an example of this approach. These authors divided verbal fluency production into two components: 1) clustering, which is the production of groups of semantically or phonemically related words (on semantic and phonemic fluency tests, respectively), and 2) switching, which is the shifting between clusters of related words (Troyer et al., 1997). These authors propose that clustering is dependent on intact temporal lobe functioning while switching relies more heavily on prefrontal lobe functioning, a distinction supported by previous lesion and neuroimaging research (Hirshorn & Thompson-Schill, 2006; Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998).

If these two components of verbal fluency production are dissociable according to anatomical and functional brain regions, then differences in clustering and switching should differentiate dementia subtypes believed to have differential effects on mesial temporal lobe and prefrontal lobe integrity. For example, dementia associated with Parkinsons's disease is presumed to impact subcortical prefrontal connectivity and therefore would be expected to show a larger effect on switching performance during verbal fluency tasks. In fact, previous researchers have reported impaired switching in contrast to average cluster size on both tasks (McDowd et al., 2011; Troster et al., 1998), and other studies have shown preserved semantic cluster size in dementia with Parkinson's disease compared to healthy controls (Epker, Lacritz, & Cullum, 1999; Troyer, Moscovitch, Winocur, Leach, & Freedman, 1998). In contrast, some studies examining groups diagnosed with Alzheimer's disease have shown both impaired clustering on phonemic and semantic fluency tasks with preserved phonemic fluency switching (Haugrud et al., 2010; Troyer et al., 1998) or preserved semantic switching (March & Pattison, 2006), while other studies have shown both impaired clustering and switching on these tasks

(Beatty, Testa, English, & Winn, 1997; Epker et al., 1999; Gomez & White, 2006; McDowd et al., 2011; Troster et al., 1998). Individuals diagnosed with aMCI or preclinical AD have shown intact phonemic fluency performance with impaired semantic cluster size production (Fagundo et al., 2008; Murphy et al., 2006) or a decline in switching with intact cluster size (Raox et al., 2008). There have only been three previous studies that compared aMCI to healthy aging on clustering and switching variables (Fagundo et al., 2008; Murphy et al., 2006; Raox et al., 2008). As well there has been no previous research comparing these variables in other dementia subtypes including VaD, DLB, and FTD. Although clustering and switching strategies might not aid in diagnosis of AD compared to healthy older adults over and above total word production on verbal fluency tasks, these variables could differentiate other dementia subtypes and provide support for which brain regions are most impaired early on in these disorders. In addition, some researchers have proposed examining other fluency measures, including switching between single words, or between a single word and a cluster word and switching between clusters of more than one word (hard and cluster switches, respectively; Abwender, Swan, Bowerman, & Connolly, 2001) and the number of novel subcategories accessed (Lanting, Haugrud, & Crossley, 2009). These variables have not been previously compared in dementia subtypes and could aid in differential diagnosis.

The objective of the current study was to investigate measures of clustering and switching performance in groups of patients diagnosed with aMCI, AD, VaD, DLB, FTD-bv and FTD-lang (combining SD and FTD-pnf groups) and in a comparison group of healthy older adults. Phonemic and semantic verbal fluency measures included total word production, number of errors, average cluster size and number of switches as defined by Troyer et al. (1997), hard and cluster switches as defined by Abwender et al. (2001), and novel and repeated clusters as defined by Lanting et al. (2009). It was hypothesized the aMCI and AD groups, compared to normal age-equivalent adults would show significantly lower semantic total word production and average cluster sizes, and intact phonemic fluency measures, due to the effects of aMCI and AD on the medial temporal lobe structures (Hodges et al. 1999; Levy & Chelune, 2007). The FTD-bv group was anticipated to show lower total word production on the phonemic task, impaired total switches, and intact cluster size scores compared to the healthy control group due to disease effects in the prefrontal lobe (Hornberger, Geng, & Hodges, 2011). The FTD-lang group was hypothesized to show the largest fluency decline compared to a healthy control group on all

fluency measures due to associated impairments in language abilities (Robillard, 2007). The VaD and DLB groups were hypothesized to show equivalent deficits on the phonemic and semantic tasks, and impaired switching and intact cluster sizes compared to healthy adults, due to disease related subcortical connectivity deficits (Brenneis et al., 2004; Price, Jefferson, Merino, Heilman, & Libon, 2005).

Methods

Participants

The healthy comparison group for the current study was recruited from the community through a mail out list provided by the Saskatchewan Council on Aging. All clinical participants were recruited from the Rural and Remote Memory Clinic in Saskatoon, Saskatchewan while participating in an interdisciplinary dementia assessment. Diagnosis was made by an interprofessional team based on the recommendations for diagnostic criteria from the CCCDTD3 (Rockwood, Bouchard, Camicioli, & Léger, 2007). Informed consent was obtained from patients and their caregivers for de-identified data to be incorporated into a larger database. To reduce the inclusion of participants with dementia due to multiple etiologies, in the current study, participants were excluded from the AD group if they had a history of stroke, evidence of vascular change on CT head scan, or significant vascular risk factors, a history of heart disease or diabetes, current high blood pressure and high cholesterol (both risk factors were required for group exclusion), or a Hachinski Ischemic Score of 5 or more (Rosen, Terry, Fund, Katzman, & Peck, 1980). Demographic data and data from a brief neuropsychological test battery are included in Table 1 for the healthy older adult (n = 26, 15 female), aMCI (n = 14, 10 female), AD (n = 22, 17 female), VaD (n = 23, 13 female), DLB (n = 11, 6 female), FTD-bv (n = 10, 6 female), and FTD-lang (n = 10, 6 female) groups. The brief neuropsychological test battery included the Wide Range Achievement Test 3rd edition (WRAT-3; Wilkinson, 1993) Reading subtest, the Mini Mental State examination (MMSE; Folstein, Folstein, & McHugh, 1975), the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, Tierney, Mohr, & Chase, 1998), and the Trail Making Test Part A and B (Trails A and B; Reitan, 1992). Univariate ANOVAs were conducted with each neuropsychological test as a dependent variable and group as the independent variable followed by Gabriel post hoc tests to compare groups.

Table 1

Demographic and Neuropsychological Battery Means (Standard Deviations) for the Healthy Older Adult, aMCI, AD, VaD, DLB, FTD-bv, and FTD-lang groups

Variable	Healthy Older Adults	aMCI	AD	VaD	DLB	FTD-bv	FTD-lang
Age	$75.8(8.2)_{a}$	$73.9(9.3)_{a}$	$73.8(8.4)_{a}$	$74.4(8.5)_{a}$	$76.5(7.8)_{a}$	$65.3(11.7)_{b}$	$70.7(11.2)_{a}$
Years of Education	$13.0(2.7)_{a}$	$10.9(3.8)_{a,b}$	$11.8(3.3)_{a,b}$	$9.7(2.7)_{b}$	$10.5(3.3)_{a,b}$	$12.3(2.1)_{a,b}$	$12.5(2.5)_{a,b}$
WRAT-3 Reading[d]	$104.2(10.7)_{a}$	$100.9(7.5)_{a,b}$	$97.6(10.6)_{a,b}$	$94.8(9.7)_{b}$	$94.8(7.5)_{a,b}$	$94.9(10.0)_{a,b}$	$89.9(11.9)_{b}$
MMSE[e]	$28.9(1.1)_{a}$	$26.9(1.9)_{a,b}$	$23.2(3.2)_{c}$	$24.4(3.8)_{b,c}$	$22.6(5.3)_{c}$	$26.1(2.3)_{a,b}$	$20.9(4.5)_{c}$
Immediate Memory[f]	$101.0(15.0)_{a}$	$75.6(13.0)_{b}$	$59.8(14.6)_{c}$	$69.1(17.9)_{b}$	$64.0(11.6)_{b}$	$74.7(11.9)_{b}$	$46.7(6.5)_{c}$
Visuospatial[f]	$100.7(16.4)_{a}$	$97.3(12.6)_{a}$	$80.7(17.5)_{b}$	$78.6(14.6)_{\mathbf{b}}$	$74.0(18.8)_{b}$	$72.4(12.9)_{b}$	$69.7(14.6)_{b}$
Language[f]	$104.4(11.0)_{a}$	$93.0(10.5)_{a,b}$	$82.5(10.4)_{b}$	$80.6(14.7)_{b}$	$79.3(15.9)_{b}$	$83.2(11.1)_{b}$	$61.1(17.2)_{c}$
Attention[f]	$101.1(17.4)_{a}$	$87.5(13.1)_{a,b}$	$74.3(17.9)_{b,c}$	$67.8(11.6)_{b,c}$	$65.7(18.0)_{b,c}$	$78.7(16.8)_{b,c}$	$58.3(12.5)_{c}$
Delayed Memory[f]	$93.3(14.2)_{a}$	$59.4(13.9)_{b}$	$52.8(12.7)_{b}$	$66.7(21.1)_{b}$	$53.6(13.4)_{b}$	$70.2(19.6)_{b}$	$55.0(16.3)_{b}$
Trails A[g]	$38.6(10.6)_{a}$	$43.4(18.6)_{a,b}$	$84.9(52.2)_{b}$	$80.0(29.1)_{b}$	$135.3(71.8)_{c}$	$79.4(49.5)_{a,b}$	$83.8(46.5)_{b,c}$
Trails B[g]	$97.5(44.6)_{a}$	$133.6(70.7)_{a,b}$	$234.3(87.1)_{c}$	$249.6(72.7)_{c}$	$272.2(63.0)_{c}$	$205.2(88.4)_{b,c}$	$290.6(29.7)_{c}$

Note. Means in each row that share subscripts do not differ significantly.

[d]The Wide Range Achievement Test 3rd edition reading subtest is a measure of single word reading with presented as scaled scores with a mean of 100 and a standard deviation of 10 (WRAT-3; Wilkinson, 1993). [e]The Mini Mental State examination is a screening measure for cognitive impairment and is scored out of a total of 30 (Folstein, Folstein, & McHugh, 1975). [f]The Repeatable Battery for the Assessment of Neuropsychological Status Index scores are scaled scores with a mean of 100 and a standard deviation of 10 (RBANS; Randolph et al., 1998). [g]Scores for the Trail Making Test Part A (Trails A) are the number of seconds taken to sequentially

join numbers in an array; scores for Part B (Trails B) are the number of seconds taken to alternate between joining numbers and letters in an array (Reitan, 1992).

Materials

As part of a comprehensive neuropsychological research battery, participants completed the Controlled Oral Word Association Test (COWAT; Benton & Hamsher, 1989) as a measure of phonemic fluency and the Animal Naming test (AN; Spreen & Strauss, 1991) as a measure of semantic fluency.

Procedures and Scoring

All neuropsychological measures were administered according to standardized instructions (COWAT, Benton & Hamsher, 1989; AN, Spreen & Strauss, 1991). The COWAT consists of three 60s trials during which participants are required to produce as many words as possible that begin with the letters "C", "F", or "L". On the Animal Naming (AN) test, participants are given 60s to produce as many animal names as possible.

Verbal fluency variables were calculated with intrusions (i.e., errors and perseverations) excluded consistent with previous research (Haugrud et al., 2010). On the phonemic task, the three trials were added together to produce a phonemic total score for each variable. Detailed scoring procedures for the calculation of clustering and switching variables have been previously reported (Abwender et al., 2001; Lanting et al., 2009; Troyer et al., 1997). Briefly, a cluster is a set of phonemically or semantically related words (on the phonemic or semantic task, respectively), while a switch is a shift between clusters. Hard switches (a switch between two single words or between a single word and clustered word), cluster switches (a switch between two groups of clustered words), number of novel clusters accessed (novel cluster), and number of previously accessed clusters returned to (repeated cluster) were also calculated.

A computer program developed to calculate clustering and switching scores was used for the current study (Haugrud et al., 2011). Use of this program has been previously supported and results in more accurate and consistent fluency scoring. This program was created with a slight modification to the original scoring measures of Troyer and colleagues (1997); on the phonemic task, only the criterion of the same first two letters was used as a cluster.

Results

Univariate ANOVAs were conducted with each fluency variable as a dependent variable and group as the independent variable with Gabriel post hoc tests used to compare groups. Partial η_p^2 values are reported as measures of effect size.

Semantic fluency

Raw scores for the semantic fluency variables can be found in Table 2 for the healthy older adult, aMCI, AD, VaD, DLB, FTD-bv, and FTD-lang groups.

Table 2

Semantic Verbal Fluency Means (Standard Deviations) for the Healthy Older Adult, aMCI, AD, VaD, DLB, FTD-bv, and FTD-lang Groups

Variable	Healthy Older Adults	aMCI	AD	VaD	DLB	FTD-bv	FTD-lang
Total Words Produced[d]	$17.9(4.6)_{a}$	$13.5(4.2)_{b}$	$9.4(2.8)_{b}$	$8.7(4.2)_{c}$	$7.8(4.8)_{c}$	$10.3(3.4)_{b}$	$6.4(3.9)_{c}$
Total Errors[e]	$1.8(2.4)_{a}$	$1.7(1.9)_{a}$	$1.9(1.3)_{a}$	$1.2(1.5)_{a}$	$0.8(1.3)_{a}$	$1.8(2.8)_{a}$	$2.4(3.0)_{a}$
Average Cluster Size[f]	$1.30(0.62)_{a}$	$0.95(0.47)_{a,b}$	$0.79(0.48)_{b}$	$1.10(0.70)_{a,b}$	$0.87(0.48)_{a,b}$	$0.70(0.35)_{a,b}$	$0.64(0.38)_{b}$
Total Switches[g]	$7.2(2.2)_{a}$	$6.1(1.9)_{a,b}$	$4.7(2.0)_{b,c}$	$3.7(2.6)_{c}$	$3.0(1.9)_{c}$	$5.2(2.2)_{a,b}$	$2.9(2.8)_{c}$
Hard Switches[h]	$4.9(2.7)_{a}$	$4.7(1.4)_{a,b}$	$3.8(2.4)_{a,b}$	$2.7(2.2)_{b}$	$1.8(1.4)_{c}$	$4.3(2.3)_{a,b}$	$2.4(1.9)_{b}$
Cluster Switches[i]	$2.4(1.4)_{a}$	$1.4(1.2)_{a,b}$	$0.9(1.1)_{b}$	$1.0(1.1)_{b}$	$1.2(1.2)_{a,b}$	$0.9(1.1)_{b}$	$0.5(0.7)_{b}$
Novel Clusters[j]	$6.6(1.8)_{a}$	$5.0(1.3)_{a,b}$	$4.6(1.3)_{b}$	$4.0(1.8)_{b}$	$3.6(1.9)_{b}$	$4.6(1.5)_{b}$	$3.4(1.7)_{b}$
Repeated Clusters[k]	$1.6(1.5)_{a,b}$	$2.1(1.3)_{a}$	$1.1(0.9)_{a,b}$	$0.7(1.1)_{b}$	$0.3(0.5)_{b}$	$1.6(1.3)_{a,b}$	$0.5(0.9)_{b}$

Note. Means in each row that share subscripts do not differ significantly.

[d]Total words produced is scored as the sum of all words on a 60 second trial minus errors and perseverations. [e]Total errors is scored as the sum of all errors and perseverations across a trial. [f]Average cluster size is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one. [g]Total switches is scored as the number of clusters of words on a 60 second trial minus one. [h]Hard switches is the sum of switches between two single words or between a single word and a clustered word on a 60 second trial. [i]Cluster switches is scored as the sum of switches between two clusters of more than one word on a 60 second trial. [j]Novel clusters is scored as the sum of new animal subcategories accessed during a 60 second trial. [k]Repeated clusters is scored as the sum of all animal subcategories an individual returns to on a 60 second trial.

Traditional scoring methods.

The main effect of group was significant for semantic fluency total words produced, $F(6,109) = 18.600$, $p < .001$, $\eta_p^2 = .506$. Consistent with the study hypothesis, all clinical groups produced fewer total words than the control group. As well, the VaD, DLB, and FTD-lang groups produced fewer words than the aMCI group. There was no group difference for number of errors, $F(6,109) = 0.833$, $p = .547$, $\eta_p^2 = .044$.

Troyer and colleagues (1997) scoring methods.

There was a significant main effect of group for average cluster size, $F(6,109) = 3.231$, $p = .006$, $\eta_p^2 = .151$. Consistent with the study hypotheses, the control group produced larger average clusters than the AD and FTD-lang groups but the average cluster size scores of the FTD-bv, VaD, and DLB groups were relatively intact. In contrast to the study hypotheses, the aMCI group also showed intact semantic average cluster size compared to the control group. The main effect of group was significant for semantic fluency total switches, $F(6,109) = 9.357$, $p < .001$, $\eta_p^2 = .340$. Consistent with the study hypotheses, the control group produced more switches than the AD, VaD, DLB, and FTD-lang groups but there was no difference observed between the control group and the aMCI or AD groups. The aMCI group also produced more switches than the VaD, DLB, and FTD-lang groups, indicating semantic fluency switches further differentiated the clinical groups, rather than simply differentiating clinical groups from the control group.

Abwender and colleagues (2001) scoring methods.

The main effect of group was significant for semantic fluency hard switches, $F(6,109) = 4.519$, $p < .001$, $\eta_p^2 = .199$. The control group produced more hard switches than the VaD, DLB, and FTD-lang groups, while the aMCI and AD groups showed no difference from the control group, consistent with the study hypotheses. Contrary to the study hypothesis, the FTD-bv group did not produce fewer hard switches than the control group. In addition, the aMCI group produced more hard switches than the DLB group, indicating hard switches can also differentiate clinical effects. As well there was a significant main effect of cluster switches, $F(6,109) = 4.113$, $p = .001$, $\eta_p^2 = .185$. The control group produced more cluster switches than the AD and FTD-lang groups.

Lanting and colleagues (2009) scoring methods.

For semantic fluency novel clusters there was a significant group effect, $F(6,109) = 8.445$, $p < .001$, $\eta_p^2 = .317$. The control group produced more novel clusters than the AD, VaD,

DLB, and FTD-lang groups. There was also a significant effect for repeated clusters, $F(6, 109) = 4.918$, $p < .001$, $\eta_p^2 = .213$, with the VaD and DLB groups producing fewer repeated clusters than the control group. As well, the aMCI group produced more repeated clusters than the VaD, DLB, and FTD-lang groups. Because novel and repeated clusters had not been previously compared in groups diagnosed with dementia, no hypotheses were proposed for these variables. However, the results indicate novel and repeated clusters show similar group difference to total switches, with the VaD, DLB, and FTD-lang groups showing the most impaired performance while the AD and aMCI groups are relatively preserved compared to the control group.

Phonemic fluency

Raw scores for the phonemic fluency variables can be found in Table 3 for the healthy older adult, aMCI, AD, VaD, DLB, FTD-bv, and FTD-lang groups.

Table 3

Phonemic Verbal Fluency Means (Standard Deviations) for the Healthy older adult, aMCI, AD, VaD, DLB, FTD-bv, and FTD-lang groups.

Variable	Healthy Older Adults	aMCI	AD	VaD	DLB	FTD-bv	FTD-lang
Total Words Produced[d]	$34.7(13.0)_{a}$	$34.0(7.8)_{a,b}$	$25.1(10.0)_{b,c}$	$18.7(9.9)_{c}$	$17.6(12.3)_{c}$	$17.9(11.1)_{c}$	$13.4(5.9)_{c}$
Total Errors[e]	$2.9(2.3)_{a}$	$1.8(2.1)_{a}$	$2.5(2.3)_{a}$	$2.1(2.4)_{a}$	$2.2(3.1)_{a}$	$4.6(3.7)_{a}$	$5.3(4.6)_{a}$
Average Cluster Size[f]	$0.50(0.27)_{a}$	$0.39(0.18)_{a}$	$0.40(0.24)_{a}$	$0.25(0.18)_{a}$	$0.34(0.34)_{a}$	$0.47(0.41)_{a}$	$0.41(0.30)_{a}$
Total Switches[g]	$20.9(6.8)_{a}$	$22.9(7.4)_{a}$	$16.3(8.1)_{a,b}$	$12.2(7.7)_{b,c}$	$10.6(8.5)_{b,c}$	$10.0(8.2)_{b,c}$	$7.1(4.3)_{c}$
Hard Switches[h]	$19.1(5.8)_{a}$	$22.2(7.5)_{a}$	$15.6(7.9)_{a,b}$	$12.1(7.7)_{b,c}$	$9.9(7.8)_{b,c}$	$9.4(7.9)_{b,c}$	$6.9(4.3)_{c}$
Cluster Switches[i]	$1.7(1.6)_{a}$	$0.7(0.9)_{a,b}$	$0.7(0.8)_{b}$	$0.1(0.3)_{b}$	$0.6(1.0)_{b}$	$0.6(1.3)_{a,b}$	$0.2(0.4)_{b}$
Novel Clusters[j]	$11.9(2.9)_{a,b}$	$13.4(2.4)_{a}$	$11.5(3.0)_{a,b,c}$	$9.4(3.3)_{b,c}$	$9.0(4.8)_{b,c}$	$8.5(3.6)_{b,c}$	$8.0(2.7)_{c}$
Repeated Clusters[k]	$11.9(4.8)_{a}$	$12.6(5.7)_{a}$	$7.8(5.7)_{a,b}$	$5.7(5.1)_{b}$	$4.2(4.6)_{b}$	$4.4(4.9)_{b}$	$2.0(2.0)_{b}$

Note. Means in each row that share subscripts do not differ significantly.

[d]Total words produced is scored as the sum of all words on a 60 second trial minus errors and perseverations; the three phonemic trials are summed to produce a total phonemic score. [e]Total errors is scored as the sum of all errors and perseverations across a trial; the three phonemic trials are summed to produce a total phonemic score. [f]Average cluster size is scored as the average of all clusters produced across a 60 second trial where clusters are scored as the number of words in a cluster minus one; the three phonemic trials are averaged to produce a total average phonemic score. [g]Total switches is scored as the number of clusters of words on a 60 second trial minus one; the three phonemic trials are summed to produce a total phonemic score. [h]Hard switches is the sum of switches between two single words or between a single word and a clustered word on a 60 second trial; the three phonemic trials are summed for a total phonemic score. [i]Cluster switches is scored as the sum of switches between two clusters of more than one word on a 60

second trial; the three phonemic trials are summed for a total phonemic score. [j]Novel clusters is scored as the sum of new phonemic subcategories accessed during a 60 second trial; the three phonemic trials are summed for a total phonemic score. [k]Repeated clusters is scored as the sum of all phonemic subcategories an individual returns to on a 60 second trial; the three phonemic trials are summed for a total phonemic score.

Traditional scoring methods.

The main effect of group was significant for phonemic fluency total words, $F(6,109) = 10.343$, $p < .001$, $\eta_p^2 = .363$. Consistent with the study hypotheses, the control group produced more words than the VaD, DLB, FTD-bv, and FTD-lang groups while the aMCI and AD groups showed relatively intact performance. In addition, the aMCI group produced more words than the VaD, DLB, FTD-fv, and FTD-lang groups, indicating group scores on phonemic total words produced also differentiated subtypes of dementia. The main effect of group was also significant for total errors, $F(6,109) = 2.722$, $p = .017$, $\eta_p^2 = .130$, with the FTD-lang group producing more errors than the aMCI and VaD groups.

Troyer and colleagues (1997) scoring methods.

Consistent with the study hypotheses, the main effect was not significant for phonemic average cluster size, $F(6,109) = 1.940$, $p = .081$, $\eta_p^2 = .096$. The main effect was significant for phonemic total switches, $F(6,109) = 9.236$, $p < .001$, $\eta_p^2 = .337$. The control group produced more switches than the VaD, DLB, FTD-bv, and FTD-lang groups, consistent with the study hypotheses. In contrast, the aMCI and AD groups showed intact scores on phonemic total switches, also consistent with the study hypotheses. The aMCI group produced more switches than the VaD, DLB, FTD-fv, and FTD-lang groups and the AD group produced more switches than the FTD-lang group. This indicates phonemic fluency total switches both differentiates healthy older adults from individuals with cognitive impairment and differentiates subtypes of dementia.

Abwender and colleagues (2001) scoring methods.

There was a significant effect for phonemic hard switches, $F(6,109) = 8.532$, $p < .001$, $\eta_p^2 = .320$ with the control group producing more hard switches than the VaD, DLB, FTD-bv, and FTD-lang groups. Additionally the aMCI group produced more hard switches than the VaD, DLB, FTD-bv, and FTD-lang groups and the AD group produced more hard switches than the FTD-lang group. Observed group difference on phonemic hard switches were consistent with the study hypotheses and are consistent with the use of the variable hard switches to differentiate subtypes of dementia. The main effect was significant for cluster switches, $F(6,109) = 5.759$, $p < .001$, $\eta_p^2 = .241$. The control group produced more cluster switches than the AD, VaD and FTD-lang groups.

Lanting and colleagues (2009) scoring methods.

The main effect of group was significant for novel clusters, $F(6,109) = 5.574$, $p < .001$, $\eta_p^2 = .235$. The only group that differed from the control group was the FTD-lang group, which produced fewer novel clusters than the control group. Although the aMCI group and the control group did not differ, the aMCI group produced more novel clusters than the VaD, DLB, FTD-bv, and FTD-lang groups. There was a significant effect of group for repeated clusters, $F(6,109) = 9.753$, $p < .001$, $\eta_p^2 = .349$. The control group produced more repeated clusters than the VaD, DLB, FTD-bv, and FTD-lang groups. The aMCI group produced more repeated clusters than the VaD, DLB, FTD-bv, and FTD-lang groups and the AD group produced more repeated clusters than the FTD-lang group. Because the variables of novel and repeated clusters were had not been previously compared in a dementia sample no hypotheses were proposed for group differences on these variables. However, observed results indicate novel and repeated clusters on phonemic fluency separate groups in a similar manner to phonemic fluency total switches. Therefore novel and repeated clusters might not provide additional information above and beyond a score of total switches for differentiating subtypes of dementia on the phonemic fluency task.

Discussion

The aim of the current study was to compare healthy older adults and individuals diagnosed with aMCI and dementia on a range of verbal fluency measures, including measures of clustering and switching. The results are consistent with previous research examining total word production in dementia. When compared to healthy older adults, the aMCI group show impaired semantic total word production with intact phonemic fluency production. The AD group was impaired on both fluency tasks compared to healthy older adults but the effect was smaller for the phonemic task. In contrast, although the FTD-bv group had impaired total word production on both tasks compared to the healthy older adult group, this effect was larger on the phonemic task. The VaD, DLB, and FTD-lang groups showed equally impaired performance on both semantic and phonemic fluency tasks. Total word production on verbal fluency tasks therefore is a sensitive measure of dementia and is useful for differentiating some dementia subtypes. Use of clustering and switching scores in the current study provided additional information on the source of fluency decline in dementia.

The aMCI group produced normal scores on measures of clustering and switching. In contrast, the AD group produced significantly lower scores on semantic fluency total switches

and average cluster size compared to the control group. Previous researchers that have shown both reduced switching and clustering in AD have concluded that these measures do not provide additional information over and above total word production (McDowd et al., 2011). However in the current study the lower semantic switching scores in the AD group were not due to simply reduced total switching, but rather was limited to a decline in cluster switching. The AD group showed lower switching between groups of semantically related words but intact switching between single words. Abwender and colleagues (2001) proposed hard switching to be dependent on processing speed while cluster switching is more dependent on intact strategic search abilities of semantic memory storage. Impaired semantic cluster size and cluster switching, with intact hard switching, observed in the AD group in this study, provides support for decreased semantic storage integrity in AD (Hodges et al. 1999; Levy & Chelune, 2007). As well, the only other clinical group that showed this pattern of impairment was the FTD-lang group which had the lowest fluency production overall and showed impairment on all clustering and switching measures. Lower average cluster size and number of cluster switches during semantic fluency, therefore, represents a unique pattern for the AD group.

Clustering and switching variables also differentiated the FTD groups. On the phonemic task, lower total word production in the FTD-bv group was associated with lower switching rates, particularly hard switching. These results are consistent with impaired prefrontal lobe functioning and impaired search and retrieval strategies in FTD-bv (Levy & Chelune, 2007). In contrast, the FTD-lang group produced the most impaired scores when compared to the other clinical groups across all measures of phonemic and semantic fluency, including producing the most errors. In a dementia subtype that primarily affects the prefrontal lobe structures at early stage (i.e., the FTD-bv group), lower fluency performance is associated with lower processing speed and mental set shifting. In a dementia subtype that affects language production at early stages (i.e., the FTD-lang group), all fluency production is lower and clustering and switching scores are correspondingly lower across all measures.

The VaD and DLB groups produced similar patterns of impaired fluency performance. Both groups produced fewer total switches compared to the healthy control group (due to fewer hard switches with intact cluster switching). This is similar to the FTD-bv group. However, the FTD-bv group showed intact semantic fluency switching. The VaD and DLB groups showed equivalently reduced hard switching on both semantic and phonemic fluency with intact cluster

sizes. This pattern is consistent with impaired diffuse subcortical and cortical connections and impaired processing speed, mental set shifting, and cognitive flexibility in VaD and DLB (Brenneis et al., 2004; Price, Jefferson, Merino, Heilman, & Libon, 2005).

In the current study, the AD participants, presumably the subgroup with impaired mesial temporal lobe structures, had impaired semantic cluster sizes and impaired cluster switching. The participants with impaired prefrontal lobe integrity, FTD-bv, showed impaired phonemic hard switching with intact cluster size scores. Individuals with more diffuse subcortical impairment, VaD and DLB, showed equivalent impairment on both fluency tasks due to reduced hard switching. Although total word production is highly sensitive to impairment in both preclinical (aMCI) and early stage dementia, clustering and switching scores provide additional information consistent with underlying structural brain impairment. The variables proposed by Troyer et al. (1997) and Abwender et al. (2001) differentiated dementia groups, providing validity for the use of these variables in this clinical population. The variables proposed by Lanting et al. (2009), however, did not reliably aid in differentiating groups. Novel clusters generation was impaired across clinical groups for semantic fluency, indicating all clinical groups had difficulty accessing semantic subcategories. For the phonemic task, novel and repeated clusters were reduced for the FTD-bv, VaD, DLB, and FTD-lang groups. Novel and repeated clusters therefore do not add additional information over and above differences in total word production for these dementia subtypes.

The current study includes normal participants, individuals thought to be in the preclinical stages of dementia (i.e., aMCI), and groups of participants from common dementia diagnostic subgroups who were all in the relatively early stages of dementia. All data were collected on the day of initial diagnosis, but clustering and switching scores were not included in the diagnostic decision making and MMSE scores are consistent with early stage dementia for all dementia subgroups. As well, the AD group was specifically selected to exclude individuals with potential multiple dementia presentation. Consequently, individuals with a history of stroke, evidence of vascular pathology on CT or significant vascular risk factors (i.e. diabetes, high blood pressure and history of potential vascular event) were excluded from the AD group. Prior researchers have noted that vascular risk factors and stroke increase the risk of developing AD and individuals with increased vascular risk factors and AD have poorer performance on measures of episodic memory compared to those without risk factors (Reitz et al., 2007; Skoog,

2004). It is possible that inconsistent results from previous studies with respect to clustering and switching in AD are due to the inclusion of individuals with both VaD and AD. In addition, previous research that has reported switching and clustering decline in AD has failed to analyze hard and cluster switching. Consequently, previous researchers have concluded that clustering and switching variables are both equally impaired in AD and therefore have little utility over the traditional measures of total words produced. In contrast to these conclusions, the results of the current study indicate that individuals in the early stages of AD are particularly vulnerable to the effects of cluster switching but perform normally on measures of hard switching. Future research should make a distinction between hard and cluster switching when analyzing dementia subgroups.

A limitation of the current study is the small sample size for the FTD-bv, FTD-lang, and DLB groups. It is important to replicate the current findings with larger samples of these clinical groups to ensure clustering and switching differences are robust. As well, the current study used only one semantic subcategory (i.e., animals). Previous research has noted that performance for both healthy individuals and individual with AD is impacted by the difficulty of subcategory used for semantic fluency (Brandt & Manning, 2009). Future work should extend this analysis to other semantic fluency tasks such as fruits and vegetables or tools to determine whether clustering and switching group differences are limited to animal naming semantic fluency. In the current study, verbal fluency data was used as part of a broader neuropsychological test battery for the purposes of dementia diagnosis. Therefore, participant fluency data contributed to initial diagnosis and then this diagnosis was used to group individuals for further analysis of the subcomponents of the fluency data. This dual, circular use of fluency data is a limitation of the current study, although as described above, the subcomponent data were not used during the clinical diagnosis.

Verbal fluency tasks are frequently used neuropsychological assessment measures for the diagnosis of dementia. Understanding what cognitive abilities and associated brain structures are required for normal performance on these tasks and how these tasks are impacted by dementia subtypes contributes to our understanding of cognitive decline in dementia. Although a score of total word production on fluency tasks is able to detect impairment, further analysis of fluency performance through clustering and switching measures is needed to achieve a detailed picture of how and why production differences occur between dementia subtypes. The results of the current

study provide evidence for the validity of measures of clustering and switching in differentiating dementia subtypes. Based on the current research it is recommended that the method of calculating average cluster size proposed by Troyer and colleagues (1997) and the methods of calculating hard and cluster switching as defined by Abwender and colleagues (2001) be used in future research examining verbal fluency and dementia. These variables best differentiated dementia subtypes from healthy aging and also were relatively easy to calculate using computerized scoring.

References

Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment, 8*, 323-336. doi: 10.1177/107319110100800308

Alzheimer Society (2010). *Rising tide: The impact of dementia on Canadian society*. Retrieved from Alzheimer Society website: http://www.alzheimer.ca/english/rising_tide/rising_tide.htm

Beatty, W. W., Testa, J. A., English, S., & Winn, P. (1997). Influences of clustering and switching on the verbal fluency performance of patients with Alzheimer's disease. *Aging, Neuropsychology, and Cognition, 4*, 273-279. doi: 10.1080/13825589708256652

Benton, A. L., & Hamsher, K. (1989). *Multilingual aphasia examination.* Iowa City, Iowa: AJA Associates.

Braaten, A. J., Parsons, T. D., McCue, R., Sellers, A., & Burns, W. J. (2006). Neurocognitive differential diagnosis of dementing diseases: Alzheimer's dementia, vascular dementia, frontotemporal dementia, and major depressive disorder. *International Journal of Neuroscience, 116*, 1271-1293. doi: 10.1080/00207450600920928

Brandt, J., & Manning, K. J. (2009). Patterns of word-list generation in mild cognitive impairment and Alzheimer's disease. *The Clinical Neuropsychologist, 23*, 870-879. doi: 10.1080/13854040802585063

Brenneis, C., Wenning, G. K., Egger, K. E., Schocke, M., Trieb, T., Seppi, K.,…Poewe, W. (2004). Basal forebrain atrophy is a distinctive pattern in dementia with Lewy bodies, 15, 1711-1714. doi: 10.1097/01.wnr.0000136736.73895.03

Canning, S. J., Leach, L., Stuss, D., Ngo, L., & Black, S. E. (2004). Diagnostic utility of abbreviated fluency measures in Alzheimer disease and vascular dementia. *Neurology, 62*, 556-562.

Chertkow, H., Nasreddine, Z., Joanette, Y., Drolet, V., Kirk, J., Massoud, F., Belleville, S., & Bergman, H. (2007). Mild cognitive impairment and cognitive impairment, no dementia: Part A, concept and diagnosis. *Alzheimer's & Dementia, 3*, 266-282.

Crossley, M., D'Arcy, C., & Rawson, N. S. B. (1997). Letter and category fluency in community-dwelling Canadian seniors: A comparison of normal participants to those

with dementia of the Alzheimer or vascular type. *Journal of Clinical and Experimental Neuropsychology, 19*, 52-62. doi: 10.1080/01688639708403836

Epker, M. O., Lacritz, L. H., & Munro Cullum, C. (1999). Comparative analysis of qualitative verbal fluency performance in normal elderly and demented populations. *Journal of Clinical and Experimental Neuropsychology, 21*, 425-434. doi: 10.1076/jcen.21.4.425.890

Fagundo, A. B., Lopez, S., Romero, M., Guarch, J., Marcos, T., & Salamero, M. (2008). Clustering and switching in semantic fluency: Predictors of the development of alzheimer's disease. *International Journal of Geriatric Psychiatry, 23*, 1007-1013. doi: 10.1002/gps.2025

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-198. doi: 10.1016/0022-3956%2875%2990026-6

Giovagnoli, A. R., Erbetta, A., Reati, F., & Bugiani, O. (2008). Differential neuropsychological patterns of frontal variant frontotemporal dementia and Alzheimer's disease in a study of diagnostic concordance. *Neuropsychologia, 46,* 1495-1504.

Gomez, R. G., & White, D. A. (2006). Using verbal fluency to detect very mild dementia of the Alzheimer type. *Archives of Clinical Neuropsychology, 21,* 771-775. doi: 10.1016/j.acn.2006.06.012

Haugrud, N., Lanting, S., & Crossley, M. (2010). The effects of age, sex and Alzheimer's disease on strategy use during verbal fluency tasks. *Aging, Neuropsychology, & Cognition, 17.* 220-239. doi: 10.1080/13825580903042700.

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia, 42*, 1212-1222. doi: 10.1016/j.neuropsychologia.2004.02.001

Hirshorn, E. A., & Thompson Schill, S. L. (2006). Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency. *Neuropsychologia, 44*, 2547-2557. doi: 10.1016/j.neuropsychologia.2006.03.035

Hodges, J. R., Patterson, K., Ward, R., Garrard, P., Bak, T., Perry, R., et al. (1999). The differentiation of semantic dementia and frontal lobe dementia (temporal and frontal

variants of frontotemporal dementia) from early Alzheimer's disease: A comparative neuropsychological study. *Neuropsychology, 13,* 31-40. doi: 10.1037/0894-4105.13.1.31

Hornberger, M., Geng, J., & Hodges, J. R. (2011). Convergent grey and white matter evidence of orbitofrontal cortex changes related to disinhibition in behavioural variant frontotemporal dementia. *Brain, 134*, 2502-2512. doi: 10.1094/brain/awr173.

Lafosse, J., Reed, B., Mungas, D., Sterling, S., Wahbeh, H., & Jagust, W. (1997). Fluency and memory differences between ischemic vascular dementia and Alzheimer's disease. *Neuropsychology, 11*, 514-522. doi: 10.1037/0894-4105.11.4.514

Lanting, S., Haugrud, N., & Crossley, M. (2009). The effects of age and sex on clustering and switching during speeded verbal fluency tasks. *Journal of the International Neuropsychological Society, 15*, 196-204. doi: 10.1017/S1355617709090237

Laws, K. R., Duncan, A., & Gale, T. M. (2010). 'Normal' semantic-phonemic fluency discrepancy in Alzheimer's disease? A meta-analytic study. *Cortex, 46*, 595-601. doi: 10.1016/j.cortex.2009.04.009

Levy, J. A., & Chelune, G. J. (2007). Cognitive-behavioral profiles of neurodegenerative dementias: Beyond Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology, 20*, 227-236. doi: 10.1177/0891988707309906

March, E. G., & Pattison, P. (2006). Semantic verbal fluency in Alzheimer's disease: Approaches beyond the traditional scoring system. *Journal of Clinical and Experimental Neuropsychology, 28*, 549-566. doi: 10.1080/13803390590949502

Marczinski, C. A., & Kertesz, A. (2006). Category and letter fluency in semantic dementia, primary progressive aphasia, and Alzheimer's disease. *Brain and Language, 97*, 258-265. doi: 10.1016/j.bandl.2005.11.001

McDowd, J., Hoffman, L., Rozek, E., Lyons, K. E., Pahwa, R., Burns, J., & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychology, 25*, 210-225. doi: 10.1037/a0021531

Mok, E. H. L., Lam, L. C. W., & Chiu, H. F. K. (2004). Category verbal fluency test performance in Chinese elderly with Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 18*, 120-124. doi: 10.1159/000079190

Murphy, K. J., Rich, J. B., & Troyer, A. K. (2006). Verbal fluency patterns in amnestic mild cognitive impairment are characteristic of Alzheimer's type dementia. *Journal of the*

International Neuropsychological Society, 12, 570-574. doi: 10.1017/S1355617706060590

Nutter-Upham, K. E., Saykin, A. J., Rabin, L. A., Roth, R. M., Wishart, H. A., Pare, N., & Flashman, L. A. (2008). Verbal fluency performance in amnestic MCI and older adults with cognitive complaints. *Archives of Clinical Neuropsychology, 23*, 229-241. doi: 10.1016/j.acn.2008.01.005

Oda, H., Yamamoto, Y., & Maeda, K. (2009). The neuropsychological profile in dementia with Lewy bodies and Alzheimer's disease. *International Journal of Geriatric Psychiatry, 24*, 125-131. doi: 10.1002/gps.2078

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine, 256*, 183-194.

Price, C. C., Jefferson, A. L., Merino, J. G., Heilman, K. M., & Libon, D. J. (2005). Subcortical vascular dementia: Integrating neuropsychology and neuroradiologic data. *Neurology, 65*, 376-382. doi: 10.1212/01.wnl.0000168877.06011.15

Ralph, M. A., Howard, D., Whitworth, A. B., Garrard, P., & Hodges, J. R. (2001). Semantic memory is impaired in both dementia with Lewy bodies and dementia of the Alzheimer's type: A comparative neuropsychological study and literature review. *Journal of Neurology, Neurosurger, & Psychiatry, 70*, 149-156. doi: 10.1136/jnnp.70.2.149

Randolph, C., Tierney, M. C., Mohr, E., & Chase, T. N. (1998). The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): Preliminary clinical validity. *Journal of Clinical and Experimental Neuropsychology, 20*, 310-319.

Raoux, N., Amieva, H., Le Goff, M., Auriacombe, S., Carcaillon, L., Letenneur, L., et al. (2008). Clustering and switching processes in semantic verbal fluency in the course of Alzheimer's disease subjects: Results from the PAQUID longitudinal study. *Cortex, 44*, 1188-1196. doi: 10.1016/j.cortex.2007.08.019

Rascovsky, K., Salmon, D. P., Hansen, L. A., Thal, L. J., & Galasko, D. (2007). Disparate letter and semantic category fluency deficits in autopsy-confirmed frontotemporal dementia and Alzheimer's disease. *Neuropsychology, 21*, 20-30. doi: 10.1037/0894-4105.21.1.20

Reitan, R. M. (1992). *Trail Making Test: manual for administration and scoring*. Arizona: Reitan Neuropsychology Laboratory.

Reitz, C., Patel, B., Tang, M. X., Manly, J., Mayeux, R., & Luchsinger, J. A. (2007). Relation between vascular risk factors and neuropsychological test performance among elderly persons with Alzheimer's disease. *Journal of the Neurological Sciences, 257*, 194-201. doi: 10.1016. j.jns.2007.01.030

Robillard, A. (2007). Clinical diagnosis of dementia. *Alzheimer's & Dementia, 3*, 292-298. doi: 10.1016/j.jalz.2007.08.002

Rockwood, K., Bouchard, R. W., Camicioli, R., & Leger, G. (2007). Toward a revision of criteria for the dementias. *Alzheimer's & Dementia, 3*, 428-440. doi: 10.1016/j.jalz.2007.07.014

Roman, G., Tatemichi, T. K., Erkinjuntti, T., Cummings, J. L., Masdeu, J. C., Garcia, J. H., … Scheinberg, P. (1993). Vascular dementia: Diagnostic criteria for research studies. Report of the NINDS-AIREN International workshop. *Neurology, 43*, 250-260.

Rosen, W. G., Terry, R. G., Fuld, P. A., Katzman, R., & Peck, A. (1979). Pathological verification of ischemic score in differentiation of dementias. *Annals of Neurology, 7*, 481-488.

Sarazin, M., Berr, C., De Rotrou, J., Fabrigoule, C., Pasquier, F., Legrain, S.,…Dubous, B. (2007). Amnestic syndrome of the medial temporal type identifies prodromal AD: A longitudinal study. *Neurology, 69*, 1859-1867.Skoog, I. (2004). Subcortical vascular dementia. The Clinical Neuropsychologist, 18, 4-5. doi: 10.1080/13854040490507109

Spreen, O., & Strauss, E. (1991). *A compendium of neuropsychological tests*. New York: Oxford University Press.

Troster, A. I., Fields, J. A., Testa, J. A., Paul, R. H., Blanco, C. R., Hames, K. A.,...Beatty, W. W. (1998). Cortical and subcortical influences on clustering and switching in the performance of verbal fluency tasks. *Neuropsychologia, 36*, 295-304. doi: 10.1016/S0028-3932%2897%2900153-X

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*, 138-146. doi: 10.1037/0894-4105.11.1.138

Troyer, A. K., Moscovitch, M., Winocur, G., Leach, L., & Freedman, M. (1998). Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *Journal of the International Neuropsychological Society, 4*, 137-143. doi: 10.1017/S1355617798001374

Troyer, A. K., Moscovitch, M., Winocur, G., Leach, L., & Freedman, M. (1998). Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *Journal of the International Neuropsychological Society*, *4*, 137-143.

Wilkinson, G. S. (1993). *Wide range achievement test – revision 3.* Wilmington, DE: Jastak Association.

General Discussion

Verbal fluency tasks have been used extensively both in research and clinical settings. The most common methods of evaluating performance on these tasks are to investigate total word production and the number of errors produced in a one minute trial. These measures are sensitive to cognitive impairment including dementia (Braaten, Parsons, McCue, Sellers, & Burns, 2006; Crossley, D'Arcy, & Rawson, 1997; Haugrud, Lanting, & Crossley, 2010; Henry, Crawford, & Phillips, 2004; Levy & Chelune, 2007; Mok, Lam, & Chiu, 2004; Rascovsky, Salmon, Hansen, Thal, & Galasko, 2007). Impaired total word production on verbal fluency tasks, however, can result due to a number of cognitive difficulties and associated areas of brain impairment (Abwender, Swan, Bowerman, & Connolly, 2001; Gierski, Peretti, & Ergis, 2007; Henry & Phillips, 2006; Marczinski & Kertesz, 2006; Mummery, Patterson, Hodges, & Wise, 1996). Consequently, while low word production could indicate the possibility of a dementia diagnosis, this scoring method does not typically allow for differential diagnosis of dementia subtypes. Troyer and colleagues (1997) proposed analyzing two subcomponents of verbal fluency production, clustering and switching. Subsequently, additional research groups described further subdividing verbal fluency production and advocated taking a more process oriented approach to test interpretation (Abwender et al., 2001; Haugrud et al., 2010; Lanting, Haugrud, & Crossley, 2009).

While clustering and switching have been compared in groups of healthy adults (Bruicki & Rocka, 2004; Haugrud et al., 2010; Lanting et al., 2009; Troyer et al., 1997; 2000), and groups diagnosed with Alzheimer's disease (Beatty, Testa, English, & Winn, 1997; Fagundo et al., 2008; Gomez & White, 2006; Haugrud et al., 2010; Troyer, Moscovitch, Winocur, Leach, & Freedman, 1998; Troster et al., 1998) and Parkinson's disease (Troster et al., 1998; Troyer et al., 1998), previous research has not compared subcomponents of verbal fluency production across other dementia subtypes. In addition, only three prior longitudinal studies of clustering and switching in individuals diagnosed with AD have been reported and these studies have produced contradictory results (Fagundo et al., 2008; Murphy, Rich, & Troyer, 2006; Raoux et al., 2008). A further limitation of previous research in this area is the difficulty and time required to hand score clustering and switching variables. The objective of the current project was to compare methods of calculating clustering and switching in a normal aging study and across dementia subtypes to determine which methods best differentiate dementia subtypes and detect age-related

effects. Study 1 involved the development of a computer scoring program to increase reliability of clustering and switching scoring. Study 2 compared clustering and switching variables across the healthy adult age span to determine which variables showed age related change and which remained age stable. The goal of study 3 was to compare healthy older adults to individuals diagnosed with AD. Study 4 compared clustering and switching variables longitudinally in an AD sample to determine how these variables are impacted by disease progression. The goal of study 5 was to determine which variables best differentiated subtypes of dementia.

Through collaboration with a graduate student in computer sciences, study 1 aimed to develop a computerized scoring program to increase reliability and efficiency of scoring clustering and switching variables on semantic and phonemic verbal fluency tasks. Hand scoring results previously published by Lanting and colleagues (2009) were compared to computerized scoring of clustering and switching for the variables proposed by Troyer and colleagues (1997), Abwender and colleagues (2001), and Lanting and colleagues (2009). This comparison showed high consistency between computer and hand scoring for phonemic fluency variables but less consistency for semantic fluency variables. When differences between hand and computer scoring were examined manually, hand scoring of semantic fluency variables showed a significant number of errors and inconsistencies, even when scorers were well trained. This indicates that computerized scoring is more accurate and consistent for measures of clustering and switching on verbal fluency tasks. In addition, the time required for scoring was significantly reduced when using the computer scoring program compared to hand scoring. Results of this study support the use of the computer scoring program and consequently this scoring program was employed for studies 2-5.

Although previous researchers have compared clustering and switching variables between young and older adults (Bruicki & Rocka, 2004; Haugrud et al., 2010; Lanting et al., 2009; Troyer et al., 1997; 2000), no previous study had compared the variables of hard and cluster switches (Abwender et al., 2001) or novel and repeated clusters (Lanting et al., 2001) in young, middle-aged, and older adults. Study 2 compared clustering and switching strategies in 90 healthy adults divided into young (20-38 yrs), middle-aged (40-63 yrs), and older (65-82 yrs) age groups. The older age group produced fewer semantic but equivalent phonemic fluency total words when compared to middle-aged and younger groups. In addition, the older age group produced fewer total switches due to fewer hard switches on both fluency tasks compared to the

middle and young age groups, but there were no age group differences for average cluster size. Results from study 2 indicate age related declines in total word production on verbal fluency tasks result from declines in switching between groups of words. These results are consistent with age related declines in processing speed and mental set shifting, and age-related stability for memory storage (Bryan & Luszcz, 2000). Study 2 results also indicate verbal fluency age group differences do not show linear age group changes. Rather verbal fluency production in the current study was equivalent between the young and middle aged groups, but the older age group showed decline in performance. The pattern of fluency production observed in healthy adults found in study 2 can be used as a comparison when investigating groups of individuals diagnosed with dementia in studies 3-5.

Study 3 compared subcomponents of verbal fluency performance (using the variables described by Abwender et al., 2001, Lanting et al., 2001, and Troyer et al., 1997) in a group diagnosed with probable Alzheimer's disease (n = 26) to a group of healthy older adults (n = 26). Total word production showed the largest group difference, especially for semantic fluency. The AD group produced fewer semantic switches (including fewer hard and cluster switches) when compared to the healthy older adult group, whereas the groups did not differ in average cluster size. The AD group also accessed fewer novel semantic subcategories on both semantic and phonemic fluency. Overall, the AD group showed impaired performance compared to the healthy older adult group on the majority of subcomponents of semantic verbal fluency but relatively intact performance on phonemic fluency. Use of subcomponent analysis of verbal fluency production does not appear to add additional diagnostic information above total word production on fluency tasks when differentiating AD from healthy aging. However, subcomponent analysis of verbal fluency might provide information on the progression of AD over time and might provide diagnostic utility in differentiating dementia subtypes. These possibilities were examined in studies 4 and 5 respectively.

Study 4 compared verbal fluency performance longitudinally over a one and two year follow-up in individuals diagnosed with probable Alzheimer's disease. Thirty-four individuals diagnosed with AD were assessed at initial diagnoses (Time 1) and at a one-year follow up (Time 2). A subsample of 19 individuals was assessed for the third time at a two-year follow up (Time 3). When all participants were included in the analysis, significant variability was observed between individuals both on the neuropsychological test battery and on measures of

clustering and switching. For example, on the Modified Mini Mental State Examination (3MS) scores ranged from 56-99 at initial assessment, indicating heterogeneity of initial symptom severity and disease stage at initial assessment. To test the hypothesis that results were confounded by a heterogeneous AD sample, the AD group was sub-dived into those with initial 3MS score above clinical cut-off of 80 and those below clinical cut-off. For the subgroup with initial 3MS scores above clinical cut-off, phonemic total word production declined from Time 1 to Time 2 and from Time 1 to Time 3 in the subsample with two year follow up data. Phonemic fluency decline over a two year follow up was due to decline in switching, particularly hard switching. As well, for the subgroup with initial 3MS score above clinical cut-off there was no observed decline on semantic fluency from Time 1 to Time 2 but semantic fluency total words produced decline over a larger follow up period (i.e. Time 1 to Time 3). The subgroup with initial 3MS scores below clinical cut-off showed significant variability in performance. It is likely that these individuals were at a more advanced stage of AD at initial assessment and therefore were already experiencing floor effects on a number of assessment measures (i.e. results were unlikely to show further decline). Results therefore suggest that phonemic fluency switching continues to decline after initial AD diagnosis in a group at preclinical or early stages of AD, consistent with progression of AD to prefrontal lobe structures. It was hypothesized that semantic fluency average cluster size would show decline over time in this study. However, at initial assessment semantic verbal fluency scores were significantly impaired compared to healthy older adults (i.e. 2 standard deviations below published norms). It is likely, therefore, that semantic fluency variables showed floor effects in this study and were unlikely to show large decline over a two year follow up period. In contrast, at initial assessment phonemic fluency performance was relatively preserved (i.e. only one standard deviation below published norms for healthy older adults) and therefore phonemic fluency decline was more readily observable. Future research should compare clustering and switching variables over a longer follow up period to determine whether phonemic fluency performance continues to decline or whether further decline on semantic average cluster size is evident at later stages of D. Alternatively, cluster size differences may only be evident or clinically useful when comparing individuals with AD to other dementia subtypes. The aim of study 5 was to correct these limitations by comparing a more homogeneous AD group to healthy older adults and individuals diagnosed with MCI, as well as to other subtypes of dementia.

In study 5 healthy older adults (n = 26) were compared to groups diagnosed with amnestic mild cognitive impairment (aMCI; n = 14), Alzheimer's disease (AD; n = 22), vascular dementia (VaD; n = 23), Lewy Body dementia (DLB; n = 11), behavioural-variant frontotemporal dementia (FTD-bv; n = 10), and language-variant frontotemporal dementia (FTD-lang; n = 10) on measures of clustering and switching (Abwender et al., 2001; Lanting et al., 2009; Troyer et al., 1997). Although total word production was a sensitive measure of dementia in this study, clustering and switching strategies differentiated dementia subtypes. The aMCI group performed normally on all fluency measures except semantic fluency total word production. In contrast to the results of study 3, average cluster size differentiated the AD group from the healthy older adult group in this study. Reduced average cluster size on semantic fluency was only seen in the AD and FTD-lang groups (with the FTD-lang group showing broad phonemic and semantic fluency deficits). The AD group showed reduced switching, but this switching reduction was due to deficits in cluster switching, or switching between groups of clustered words. This indicates impaired average cluster size and cluster switches with intact hard switching on semantic fluency is a pattern of fluency deficit for AD, when the AD group is a homogeneous group where individuals with a history of stroke or other vascular risk factors are excluded. The DLB and VaD groups showed a pattern of fluency impairment characterized by equivalently impaired phonemic and semantic total word production due to reduced switching (particularly hard switching) and reduced novel cluster generation, which is consistent with impaired processing speed and set shifting in these disorders (Brenneis et al., 2004; Price, Jefferson, Merino, Heilman, & Libon, 2005). The FTD-bv group showed a pattern of fluency impairment characterized by larger impairment on phonemic than semantic total word production due to reduced hard switching with intact cluster size production. This FTD-bv pattern is consistent with impaired complex strategic search processing in FTD-bv (Braaten et al., 2006; Wittenberg et al., 2008). The FTD-lang group was impaired across both phonemic and semantic fluency, on both measures of average cluster size and switching, consistent with the profound language impairment in this disorder (Giovagnoli, Erbetta, Reati, & Bugiani, 2008; Levy & Chelune, 2007; Wittenberg et al., 2008).

The three primary objectives of this research program were: 1) to develop a computerized scoring program to more reliably calculate measures of clustering and switching; 2) to determine which subcomponents of verbal fluency production are sensitive to age effects and which

components remain age stable; and, 3) to determine how subcomponents of verbal fluency are impacted by the progression of dementia in AD, and which subcomponents can be used to differentiate dementia subtypes. The results of study 1 provided support for the reliability and efficacy of the computerized scoring program. The program developed proved to be more reliable and faster than hand scoring of clustering and switching variables. Study 2 addressed a limitation of previous research that clustering and switching variables had not been compared across the adult lifespan. This study showed that total word production declines with increasing age group due to the reduced ability of older adults to rapidly shift between subcategories (switching), with the size of subcategories (clustering) remaining relatively age stable. Study 2 also showed that age related effects on verbal fluency are not linear across the adult lifespan but rather fluency begins to decline at approximately age 65 for both semantic and phonemic fluency.

Studies 3-5 compared verbal fluency variables in individuals diagnosed with dementia. Although the results of study 3 appeared to be inconsistent with respect to expected semantic memory storage decline in AD, when the AD group was a more homogeneous sample which excluded potential vascular comorbidity in study 5 a distinctive pattern of impairment on subcomponents of verbal fluency emerged for AD. As well, the results of study 5 showed differential patterns of impairment on verbal fluency for aMCI, FTD-bv, FTD-lang, and VaD and DLB. Study 4 provided initial evidence for the use of clustering and switching subcomponents of verbal fluency to investigate dementia-related decline over time using a subgroup of individuals diagnosed with AD. Taken together these studies support the use of measures of clustering and switching for differentiating subtypes of dementia.

For both healthy adults and individuals diagnosed with dementia, the current research showed that not all variables that have been previously proposed to assess subcomponents of clustering and switching have utility in verbal fluency analysis. The variables of total word production and number of errors that are typically used in experimental and clinical settings showed large effect sizes in both healthy aging and dementia comparisons. Therefore the use of these variables is recommended in future research. For the variables proposed by Troyer and colleagues (1997), average cluster size on semantic fluency provided important diagnostic information when differentiating dementia subtypes. However, average cluster size on phonemic fluency did not prove to be helpful in differentiating dementia groups. Therefore, for dementia

research and clinical purposes, average cluster size is only recommended for semantic verbal fluency comparisons. The variables of hard and cluster switches proposed by Abwender and colleagues (2001) differentiated dementia subtypes and showed differential impact in healthy aging. These variables appear to provide more clinical information than the more general variable of total switches proposed by Troyer and colleagues (2001). Consequently it is recommended that future research and clinical work divide total switches into hard (i.e. switching between two single words or between a single word and a clustered word) and cluster switches (i.e. switching between two groups of clustered words) when analyzing the performance of individuals diagnosed with dementia. The variables proposed by Lanting and colleagues (2009; novel and repeated clusters) did not appear to add additional information above total switching on verbal fluency tasks in this research. It is possible that these variables could provide more clinical utility with other disorders or types of brain injury and this could be explored in future research. Taken together, this research supports the use of total word production, total errors, hard switches, cluster switches, and semantic fluency average cluster size in contributing to the understanding of verbal fluency changes in healthy aging and dementia.

A limitation of this body of research is that the clustering and switching variables used are highly correlated. In addition, some variables (e.g., phonemic fluency average cluster size), produce very small values and can have large variability among groups of individuals. This can potentially impact statistical analysis and therefore the reliability of results and their interpretation. Given that the aim of the current body of research was to identify which fluency variables most reliably differentiate healthy aging and dementia, this amount of potential variability was expected in statistical analysis. Future research should reduce the number of variables analyzed to the subgroup recommended by this project (total words, errors, hard switches, cluster switches, and semantic fluency average cluster size). This will reduce potential for overlap between variables and reduce the impact of high correlations between multiple independent measures.

A second limitation of the current body of research lies in the use of verbal fluency measures for dementia diagnosis. For studies 3, 4, and 5, dementia diagnosis was made by an interprofessional team and total scores on verbal fluency tasks were part of a larger neuropsychological battery that contributed to diagnosis. This circular use of fluency variables is a potential confound of the current research. It is possible that performance on fluency tasks

impacted initial diagnosis and therefore groups were potentially pre-categorized by fluency patterns. However, use of clustering and switching strategies to compare verbal fluency tasks is a relatively recent development and not commonly used in clinical settings. As well, the interdisciplinary diagnosis incorporates a large amount of test-related, self-report, functional report, and clinical information from multiple health professionals. Diagnosis is never made based on one source of information but rather all clinical information is incorporated into diagnosis. Therefore, while it is possible that total word production on fluency tasks influenced initial diagnosis, it is less likely that clustering and switching would have played a role in this diagnosis.

The current research was the first to compare measures of clustering and switching across subtypes of dementia beyond AD or Parkinson's disease. In addition, longitudinal research with clustering and switching variables has been restricted to individuals diagnosed with AD. Future research should replicate the current study comparing subcomponents of verbal fluency across dementia subtypes. In addition, future research should aim to extend the results of study 4 and compare other subtypes of dementia longitudinally to determine how disease progression can impact clustering and switching variables. Clustering and switching are only two components of verbal fluency production that could be compared between healthy individuals and individuals diagnosed with dementia. Future research should also examine frequencies of exemplar generation and how the frequency of words generated relates to clustering and switching scores for normal adults of all ages and for dementia subtypes across time. For example, future research could compare the content of word clusters in healthy individuals and individuals diagnosed with dementia, even if average cluster size shows no group differences.

References

Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment, 8*, 323-336.

Alzheimer Society (2010). *Rising tide: The impact of dementia on Canadian society*. Retrieved from Alzheimer Society website: http://www.alzheimer.ca/english/rising_tide/rising_tide.htm

Beatty, W. W., Testa, J. A., English, S., & Winn, P. (1997). Influences of clustering and switching on the verbal fluency performance of patients with Alzheimer's disease. *Aging, Neuropsychology, and Cognition, 4*, 273-279.

Braaten, A. J., Parsons, T. D., McCue, R., Sellers, A., & Burns, W. J. (2006). Neurocognitive differential diagnosis of dementing diseases: Alzheimer's dementia, vascular dementia, frontotemporal dementia, and major depressive disorder. *International Journal of Neuroscience, 116*, 1271-1293. doi: 10.1080/00207450600920928

Brenneis, C., Wenning, G. K., Egger, K. E., Schocke, M., Trieb, T., Seppi, K.,…Poewe, W. (2004). Basal forebrain atrophy is a distinctive pattern in dementia with Lewy bodies, 15, 1711-1714. doi: 10.1097/01.wnr.0000136736.73895.03

Bruandet, A., Richard, F., Bombois, S., Maurage, C. A., Deramecourt, V., Lebert, F., Amouyel, P., & Pasquier, F. (2009). Alzheimer disease with cerebrovascular disease and vascular dementia: Clinical features and course compared with Alzheimer disease. *Journal of Neurology, Neurosurgery, & Psychiatry, 80*, 133-140. doi: 10.1136/jnnp.2007.137851

Brucki, S. M. D., & Rocha, M. S. G. (2004). Category fluency test: Effects of age, gender and education on total scores, clustering and switching in Brazilian Portuguese-speaking subjects. *Brazilian Journal of Medical and Biological Research, 37*, 1771-1777.

Bryan, J., & Luszcz, M. A. (2000). Measurement of executive function: Consideration for detecting adult age differences. *Journal of Clinical and Experimental Neuropsychology, 22*, 40-55. doi: 10.1076/1380-3395%28200002%2922:1;1-8;FT040

Crossley, M., D'Arcy, C., & Rawson, N. (1997). Letter and category fluency in community-dwelling Canadian seniors: A comparison of normal participants to those with dementia of the Alzheimer or vascular type. *Journal of Clinical and Experimental Neuropsychology, 19*, 52-62. doi: 10.1080/01688639708403836

Fagundo, A. B., Lopez, S., Romero, M., Guarch, J., Marcos, T., & Salamero, M. (2008). Clustering and switching in semantic fluency: Predictors of the development of alzheimer's disease. *International Journal of Geriatric Psychiatry, 23*, 1007-1013. doi: 10.1002/gps.2025

Gierski, F., Peretti, C. S., & Ergis, A. M. (2007). Effects of the dopamine agonist piribedil on prefrontal temporal cortical network function in normal aging as assessed by verbal fluency. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 31*, 262-268. doi: 10.1016/j.pnpbp.2006.06.017

Giovagnoli, A. R., Erbetta, A., Reati, F., & Bugiani, O. (2008). Differential neuropsychological patterns of frontal variant frontotemporal dementia and Alzheimer's disease in a study of diagnostic concordance. *Neuropsychologia, 46*, 1495-1504.

Gomez, R. G., & White, D. A. (2006). Using verbal fluency to detect very mild dementia of the Alzheimer type. *Archives of Clinical Neuropsychology, 21*, 771-775. doi: 10.1016/j.acn.2006.06.012

Haugrud, N., Lanting, S., & Crossley, M. (2010). The effects of age, sex and Alzheimer's disease on strategy use during verbal fluency tasks. *Aging, Neuropsychology, & Cognition, 17.* 220-239. doi: 10.1080/13825580903042700

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia, 42*, 1212-1222. doi: 10.1016/j.neuropsychologia.2004.02.001

Henry, J. D., & Phillips, L. H. (2006). Covariates of production and perseveration on tests of phonemic, semantic and alternating fluency in normal aging. *Aging, Neuropsychology, and Cognition, 13*, 529-551. doi: 10.1080/138255890969537

Lanting, S., Haugrud, N., & Crossley, M. (2009). The effects of age and sex on clustering and switching during speeded verbal fluency tasks. *Journal of the International Neuropsychological Society, 15*, 196-204. doi: 10.1017/S1355617709090237

Levy, J. A., & Chelune, G. J. (2007). Cognitive-behavioral profiles of neurodegenerative dementias: Beyond Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology, 20*, 227-236. doi: 10.1177/0891988707309906

Marczinski, C. A., & Kertesz, A. (2006). Category and letter fluency in semantic dementia, primary progressive aphasia, and Alzheimer's disease. *Brain and Language, 97,* 258-265. doi: 10.1016/j.bandl.2005.11.001

Mok, E. H. L., Lam, L. C. W., & Chiu, H. F. K. (2004). Category verbal fluency test performance in Chinese elderly with Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 18,* 120-124. doi: 10.1159/000079190

Mummery, C., Patterson, K., Hodges, J., & Wise, R. (1996). Generating "tiger" as an animal name or a word beginning with T: Differences in brain activation. *Proceedings: Biological Sciences, 263,* 989-995. Retrieved from http://www.jstor.org.cyber.usask.ca/stable/50587

Murphy, K. J., Rich, J. B., & Troyer, A. K. (2006). Verbal fluency patterns in amnestic mild cognitive impairment are characteristic of Alzheimer's type dementia. *Journal of the International Neuropsychological Society, 12,* 570-574. doi: 10.1017/S1355617706060590

Onyike, C. U. (2006). Cerebrovascular disease and dementia. *International Review of Psychiatry, 18,* 423-431. doi: 10.1080/09540260600935421

Price, C. C., Jefferson, A. L., Merino, J. G., Heilman, K. M., & Libon, D. J. (2005). Subcortical vascular dementia: Integrating neuropsychology and neuroradiologic data. *Neurology, 65,* 376-382. doi: 10.1212/01.wnl.0000168877.06011.15

Raoux, N., Amieva, H., Le Goff, M., Auriacombe, S., Carcaillon, L., Letenneur, L., & Dartigues, J. F. (2008). Clustering and switching processes in semantic verbal fluency in the course of Alzheimer's disease subjects: Results from the PAQUID longitudinal study. *Cortex, 44,* 1188-1196. doi: 10.1016/j.cortex.2007.08.019

Rascovsky, K., Salmon, D. P., Hansen, L. A., Thal, L. J., & Galasko, D. (2007). Disparate letter and semantic category fluency deficits in autopsy-confirmed frontotemporal dementia and Alzheimer's disease. *Neuropsychology, 21,* 20-30. doi: 10.1037/0894-4105.21.1.20

Troster, A. I., Fields, J. A., Testa, J. A., Paul, R. H., Blanco, C. R., Hames, K. A.,...Beatty, W. W. (1998). Cortical and subcortical influences on clustering and switching in the performance of verbal fluency tasks. *Neuropsychologia, 36,* 295-304. doi: 10.1016/S0028-3932%2897%2900153-X

Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology, 22*, 370-378.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*, 138-146. doi: 10.1037/0894-4105.11.1.138

Troyer, A. K., Moscovitch, M, Winocur, G., Leach, L., & Freedman, M. (1998). Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *Journal of the International Neuropsychological Society, 4*, 137-143.

Wittenberg, D., Possin, K. L., Rascovsky, K., Rankin, K. P., Miller, B. L., & Kramer, J. H. (2008). The early neuropsychological and behavioral characteristics of frontotemporal dementia. *Neuropsychology Review, 18*, 91-102. doi: 10.1007/s11065-008-9056-z

Appendix A: Consent Form Study 1

UNIVERSITY OF SASKATCHEWAN

Department of Psychology Aging Study Project

CONSENT FORM

Researchers: Margaret Crossley, Ph. D., and
Megan O'Connell

Title:

The Effects of Normal Aging on Working Memory Capacity.

Objective:

This study is an investigation of age-related changes in the ability to pay attention to two tasks at the same time. It was designed to clarify how task difficulty or familiarity affects the ability to simultaneously perform two activities. This study will also provide information about age-based changes in memory and language.

Procedure:

Volunteers will be asked to perform a variety of activities. Some of these activities involve tests of memory and language ability. Other activities include three sets of combined tasks; finger-tapping and reading, finger-tapping and speaking, following a maze and counting. Information about individual performance will be answered by the researcher whenever possible.

The procedure will take approximately two and a half to three hours to complete and will include a rest period. There are no known risks associated with this research.

I understand that this research has been approved the University Advisory Committee on Ethics in Human Experimentation. If I have any questions, complaints, or concerns I may contact Margaret Crossley at 966-5925 or Megan O'Connell at 249-5046.

I, _____ of _____, have read the above protocol and agree to participate. The procedure and its possible risks have been explained to me and I understand them. I understand that I am free to withdraw from this study at any time without penalty of any type. I understand that although the data from this study may be published in an Honours thesis, only aggregate data will be used and that my identity will be kept confidential. I also understand that all data will be kept on file for a period of five years in accordance with the University of Saskatchewan guidelines.

_____ _____ _____

(Signature of Volunteer) (Date) (Signature of Researcher)

Appendix B: Consent Form Study 2 and 3

UNIVERSITY OF SASKATCHEWAN

Department of Psychology

CONSENT FORM

Researchers: Michelle Shaw, M. A.
Margaret Crossley, Ph. D., Supervisor

Title:

The Effects of Practice on Memory

Objective:

This study is an investigation of the effects of practice on the ability to remember words and pictures. The study will also provide information about individual differences in memory and language skills.

Procedure:

You will be asked to name pictures, and to read words, and to identify similarities among pictures and words. You will complete additional language and memory tasks and will be asked to provide information about your health and lifestyle. Any questions you may have about the study will be answered by the researcher whenever possible.

All information will be treated in a confidential manner and will be safely stored at the University of Saskatchewan under the protection of Dr. Margaret Crossley for at least five years. The information collected in this study will be published in a dissertation and may be summarized in journal articles and/or professional conference presentations. At all times, only group data will be reported; individual participants will not be identified. In addition, a general written summary of the group findings from this study will be sent to you.

Participation in this study is voluntary, and your decision to participate will not impact on any clinical services that would otherwise be available to you (e.g., assessments, treatments, etc.).

This procedure will take approximately one and a half hours to complete and will include a rest period. There are no known risks associated with this research.

If I have any questions or concerns I may contact Margaret Crossley or Michelle Shaw at 966-5925.

I, _____ of _____, have read the above protocol and agree to participate. The procedure and its possible risks have been explained to me and I understand them. I acknowledge receiving a copy of this form for my own records.

I understand that I am free to withdraw from this study at any time without penalty or loss of services. Should I decide to withdraw from the study, any information I have already provided will not be included in the analyses and will be destroyed.

_____ _____ _____

(Signature of Participant) (Date) (Signature of Researcher)

Appendix C: Clinical Consent Form Studies 4 and 5

CONSENT FORM
Rural and Remote Memory Clinic and Aging Research Centre

*You are invited to participate in a study on **Neuropsychological Abilities in Older Adults with Memory Problems**. Please read this form carefully, and feel free to ask any questions you may have.*

Researchers: Margaret Crossley, Ph.D., Registered Doctoral Psychologist, Department of Psychology, University of Saskatchewan, TEL: (306) 966-5923.

Purpose and Procedure: This is a teaching and research clinic and, with your permission, a summary of your assessment materials may be stored in a computer database and used for future teaching and research purposes. All information used for research will be anonymous and not be associated with your name or any other identifiable information.

Potential Risks: There are no known health risks associated with this assessment, but you may find that the clinical neuropsychological assessment includes some tasks that are hard to complete, and you may feel mildly frustrated and/or tired.

Potential Benefits: Your assessment will help us to better understand both your cognitive strengths and weaknesses. In particular, this will help us to identify and understand any memory difficulties that you may be experiencing, as well as your skills in other areas such as language, attention, and organizational abilities. If you allow us to include your assessment material in the neuropsychological database, this information, along with information from other patients, will contribute to our understanding of cognitive functioning in older adults with memory concerns.

Confidentiality and Storage of Data: All information provided by you for this project is confidential and will only be shared with members of the project team. The research database will contain no identifying information and will be kept on a computer and back-up storage device in a secure office in the Rural and Remote Memory Clinic under the authority of Dr. Margaret Crossley at the University of Saskatchewan. The information collected through assessment and contained in the research database may be published in journal articles and/or professional conference presentations, and/or summarized for teaching and public education purposes. At all times, only aggregate results will be reported; **your name will never appear with the results. Signed consent forms will be stored separately from the research materials in locked files.**

Participation is Voluntary: You may withdraw from the project for any reason, at any time, without penalty of any sort and without losing access to the services available through the Memory Clinic. If you choose to withdraw your consent to have your clinical assessment materials included in an anonymous research database, any information that you have contributed to the database will not be used and will be destroyed. Your decision to participate in this project will in no way impact the services you will receive as a client at the Rural and Remote Memory Clinic. Should you choose to participate, your data will be used in future research studies that will help us understand both healthy aging and dementia. Should you decline to participate, your assessment materials will only be used for clinical purposes to aid in understanding your current symptoms.

Questions: If you have any questions concerning the project, please feel free to ask at any point; you are also free to contact the researcher (Margaret Crossley) at the number given below. This project has been approved on ethical grounds by the University of Saskatchewan Behavioral Research Ethics Board. Any questions regarding your rights as a participant may be addressed to that committee through the Office of Research Services (collect at 306-966-2084).

CONSENT TO PARTICIPATE:

I, _____, have read and understood the description provided above; I have been given a chance to ask questions and my questions have been answered satisfactorily. I consent to participate in the study described above, understanding that I may withdraw this consent at any time. A copy of this consent form has been given to me for my records.

Participant Signature: _____ Phone #:_____
Caregiver Signature:_____
Investigator Signature: _____
Date: _____

<u>Principal Investigator:</u>

Margaret Crossley, Associate Professor, Department of Psychology, University of Saskatchewan. TEL (306) 966-5925 (call collect)

Appendix D: Healthy Aging Consent Form Study 5

CONSENT FORM
Rural and Remote Memory Clinic and Aging Research Centre

You are invited to participate in a study on **Attention Skills and Walking Ability in Older Adults with Memory Problems.** Please read this form carefully, and feel free to ask any questions you may have.

Researchers: Patrick Corney, B. A., Shawnda Lanting, B. A., and Jocelyn Poock, B. A., Department of Psychology, University of Saskatchewan, TEL: (306) 966-5925.
Research Assistant: Jocelyn L. Poock, B. A., Department of Psychology, University of Saskatchewan, TEL: (306) 664-6658.
Supervisor: Margaret Crossley, Ph.D., Department of Psychology, University of Saskatchewan, TEL: (306) (66-5923.

Purpose and Procedure: The purpose of this study is to learn how memory changes affect our ability to pay attention and to perform two activities at the same time. You will be asked to complete tasks that require attention and concentration, and to divide your attention between talking and walking. The study will take approximately one hour to complete, including a brief rest period.

Potential Risks: There are no known serious health risks that will result from taking part in this study, but you may find some of the tasks hard to complete, or feel mildly frustrated and/or tired. All walking includes a very slight risk of falling, but we will take every precaution to prevent falls, including while you are dividing your attention between walking and talking.

Potential Benefits: By taking part in this study you will help us to better understand how attention and memory abilities are related. We hope that our findings will contribute to the development of new methods of identifying people who are experiencing changes in memory and attention skills. The results of this study may also contribute to the development of fall prevention strategies for older adults.

Confidentiality and Storage of Data: All information provided by you for this project is confidential and will only be shared with members of the project team. The data will contain no identifying information, and will be stored separately from the consent forms in a secure office assigned to Dr. Margaret Crossley at the University of Saskatchewan. The information collected in this study may be published as part of Patrick Corney's or Shawnda Lanting's doctoral dissertations and may be presented in journal articles and/or professional conference presentations. At all times, only group results will be reported; your name will never appear with the results.

Participation is Voluntary: You may withdraw from the project for any reason, at any time, without penalty of any sort and without losing access to the services available through the Memory Clinic. If you choose to withdraw from the project, any information that you have contributed will not be used and will be destroyed.

Questions: If you have any questions concerning the project, please feel free to ask at any point; you are also free to contact the researcher (Margaret Crossley) at the number given below. This project has been approved on ethical grounds by the University of Saskatchewan Behavioral Research Ethics Board. Any questions regarding your rights as a participant may be addressed to that committee through the Office of Research Services (collect at 306-966-2084).

CONSENT TO PARTICIPATE:

I, _____, have read and understood the description provided above; I have been given a chance to ask questions and my questions have been answered satisfactorily. I consent to participate in the study described above, understanding that I may withdraw this consent at any time. A copy of this consent form has been given to me for my records.

Participant Signature: _____ Phone #:_____
Caregiver Signature:_____
Investigator Signature:_____
Date:_____

Principal Investigator:

Margaret Crossley, Associate Professor, Department of Psychology, University of Saskatchewan. TEL (306) 966-5925 (call collect)